M

Journal of Experimental Psychology

THE JOHN CRERAR LIBRARY
APR 29 1953

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *Human Resources Research Center*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

The Kinetic Depth Effect: H. WALLACH AND D. N. O'CONNELL 205

Sustained Behavior under Delayed Reinforcement: C. B. FERSTER 218

The Influence of Experimentally Induced Failure on the Retention of Material Acquired through Set and Incidental Learning: M. ABORN 225

Reminiscence Effects in Paired-Associate Learning: D. A. RILEY 232

A Test of the Mechanisms of Learning Proposed by Hull and Guthrie: D. ZEAMAN AND L. RADNER 239

Retention of Serial Nonsense Syllables As a Function of Rest-Interval Responding Rate and Meaningfulness: E. J. ARCHER 245

Studies of Distributed Practice: X. The Influence of Intralist Similarity on Learning and Retention of Serial Adjective Lists: B. J. UNDERWOOD 253

Performance of a Motor Task As a Function of Interpolation of Varying Lengths of Rest at Different Points in Acquisition: E. B. NORRIS 260

The Relation of Conditioned Response Strength to Anxiety in Normal, Neurotic, and Psychotic Subjects: K. W. SPENCE AND J. A. TAYLOR 265

American Psychological Association

Vol. 45 No. 4 April 1953

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas. Priority in publication is given to articles whose authors assume the full cost of publication.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 45, NO. 4 APRIL, 1953

THE KINETIC DEPTH EFFECT

HANS WALLACH AND D. N. O'CONNELL[1]

Swarthmore College

The problem of how three-dimensional form is perceived in spite of the fact that pertinent stimulation consists only in two-dimensional retinal images has been only partly solved. Much is known about the impressive effectiveness of binocular disparity. However, the excellent perception of three-dimensional form in monocular vision has remained essentially unexplained.

It has been proposed that some patterns of stimulation on the retina give rise to three-dimensional experiences, because visual processes differ in the spontaneous organization that results from certain properties of the retinal pattern. Rules of organization are supposed to exist according to which most retinal projections of three-dimensional forms happen to produce three-dimensional percepts and most retinal images of flat forms lead to flat forms in experience also. This view has been held mainly by gestalt psychologists.

Another approach to this problem maintains that the projected stimulus patterns are interpreted on the basis of previous experience, either visual or kinesthetic. However, such empiricistic assumptions do not explain much, unless it is made clear how those previous experiences come about whose influence is supposed to account for the current perception. Kinesthetic form perception itself presents far more complex problems than does three-dimensionality in vision, and recourse to kinesthesis appears at present quite futile. Retinal disparity can of course account for that previous visual experience insofar as *Ss* with binocular vision are concerned. Whether in the absence of binocular vision, e.g., in congenitally monocular *Ss*, head movement parallex can fully play the role of binocular parallex appears doubtful. It seems that all we have left to account for an original experience of three-dimensional form in monocular *Ss* is the assumption that certain patterns of retinal stimulation will naturally produce experience of solid form. However, once it is assumed that perception of three-dimensional form can follow directly from retinal stimulation because of spontaneous organization of visual processes alone, there is no point in postulating an influence of past experience.

Unfortunately it appears that no one has succeeded in formulating rules of

[1] Most of the work reported in this and the following paper was done while the senior author was holder of a John Simon Guggenheim Memorial fellowship.

spontaneous organization adequate to predict which pattern of retinal stimulation will lead to perceived flat figures and which one will produce three-dimensional forms. We have made a vain attempt of our own and have become convinced that the three-dimensional forms perceived in perspective drawings, photographs, etc. are indeed a matter of previous experience.[2] In this situation the search for a visual process which can account for an original perception of three-dimensional form in monocular vision becomes imperative. Such a process will be described in this paper.

When one moves about, the retinal image of solid objects lying to the side of one's path not only expands but also distorts, because the objects are seen successively from different directions. More specifically, a retinal image distorts under these conditions as if the corresponding object were rotated through a certain angle in front of the eye of a stationary observer.

This fact suggests a simple technique for the investigation of the perceptual results of these distortions. An object is placed between a punctiform light source and a translucent screen and is rotated or turned back and forth. Its shadow is observed from the other side of the screen. The shadow-casting object is placed as close to the screen as possible, whereas the distance between the light source and the object is made large. Owing to this arrangement isometric projection is closely approximated. The shadows of a great number of three-dimensional forms, solid or wire-edged, will be perceived as three-dimensional under these circumstances. The shadows of some forms will look three-dimensional *only* in such a moving presentation; that is, in none of the positions through which such a form passes during rotation will it cast a stationary shadow which looks three-dimensional. With such forms one can study this effect in isolation. It will be referred to by the term "kinetic depth effect." It is important because it answers our problem: It appears that the kinetic depth effect can cause a genuine perception of three-dimensional form in a monocular *S* whenever he moves and keeps looking at an object which does not lie directly in his path.

[2] Evidence supporting this point of view will be reported in a subsequent paper.

Similar set-ups have been used before by Miles (3) and Metzger (2). Miles presented to *S*s the shadow of a two-bladed fan wheel in rotation, the shaft of which was parallel to the screen on which the shadow was formed. His *S*s reported a large number of different motion patterns, most of them involving depth, but no attempt was made to find out whether the object that was seen in motion was a good representation of the fan wheel whose shadow was presented. Metzger also was not primarily concerned with the perception of three-dimensional form which such a set-up can yield. In fact, he investigated with great thoroughness the effects of a very special kind of arrangement which is not favorable to the emergence of stable three-dimensional forms. He had arrangements of vertical rods rotate about a vertical axis and showed their shadows in a low oblong aperture which hid the ends of the rod shadows from view. As in Miles's experiment the motion patterns seen by a given *S* during an extended period of inspection were very changeable. Moreover, naive *S*s mostly differ among each other as to whether the first movement process which they perceive is in three dimensions or takes place in a plane. Suggestion has a strong influence, in the course of longer observation as well as initially. We shall see below that the patterns of line shadows which Metzger presented to his *S*s do not contain the condition essential for the kinetic depth effect.

The depth observed in Lissajous figures (1, 4) is the result of complex effects and will be discussed in a later paper.

Experiments to Define Conditions of Kinetic Depth Perception

Experiment 1: Rotation of solids.— One of the figures which we presented to many *S*s was a solid block in the

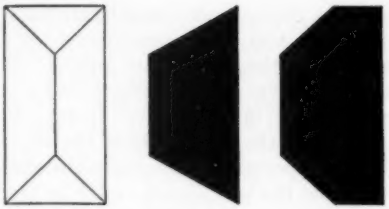


FIG. 1. Solid form (A) used in Exp. 1, and two samples (B, C) of its shadows

shape of a roof with sloping gables, as shown in Fig. 1A. This was rotated continually about its longest axis. Figures 1B and 1C show two of the forms of the shadow during rotation. When any one of the three figures is shown stationary to a naive *S*, it is described as a two-dimensional figure. When the solid is slowly rotated so that the shadow undergoes continuous deformation, *S* sees a three-dimensional solid in rotation. The direction of the seen rotation may or may not coincide with that of the object behind the screen, and this is in agreement with the conditions of stimulation which give no clue of the object's real direction of rotation. Which one of the two directions of rotation is seen seems entirely a matter of chance, and with prolonged observation *S* usually experiences a number of spontaneous reversals of the direction of rotation. Occasionally an *S* sees for long periods reversals after each rotation of 180°.

In experiments like this the impression of three-dimensional form is so natural that many *S*s who are not psychologists are not astonished by their observations. They correctly assume that behind the screen is just such an object as they see. Only after reversals of the direction of rotation have occurred do they begin to wonder.

Where the kinetic depth effect takes place, a rigid three-dimensional form in rotation is seen instead of the distorted two-dimensional figure given on the shadow screen. The distorting two-dimensional shape may occasionally be *seen* after prolonged exposure of the same kinetic presentation, or when one looks from a sharply oblique direction at the screen. The *S*'s experience of a continuously flowing form seems to him abnormal or unusual although this is exactly what is given on the retina.

The essential difference between this two-dimensional experience and the three-dimensional form which is usually seen seems to be that the latter is unchanging and rigid instead of ever changing and flowing. The changes in the shape of the retinal image are accounted for by a perceived rotation of the three-dimensional object, whereas in the two-dimensional process the seen movement distorts the form itself. As far as we can see, the distortions of the perceived two-dimensional form agree closely with the changes of the retinal image. But the perceived three-dimensional form is not determined merely by what is presented on the retina at a given moment. A single one of the projections of the shadow-casting object which make up the changing shapes of the shadow does not look three-dimensional. Only the sequence of changing shapes gives rise to the seen three-dimensional form. In other words, a single projection causes a three-dimensional form to be seen only because it was preceded by a number of other projections. By itself, it simply does not convey enough data about the three-dimensional form which it represents. The seen three-dimensional form is richer in structure than any single projection and it is built up in a temporally extended process. The individual retinal image determines only which aspect of the turning three-dimensional form appears to be given at the moment. Thus, perceptual experience far surpasses what one should expect of a process determined by momentary stimulation and seems to a higher degree a product of the immediate past than of stimulation occurring at a given moment.

Experiment 2: Partial rotation of wire figures.—Experiments of the kind just reported differ in two ways from the realistic situations in which the kinetic depth effect occurs. In the first place, a shadow-casting object is put through a complete rotation and produces a pattern of distortion on the screen which corresponds to that obtained when under natural viewing conditions an *S* moves completely around an object, which is hardly ever done. In the second place, in the shadow of a solid object an edge of the three-dimensional form is visible only

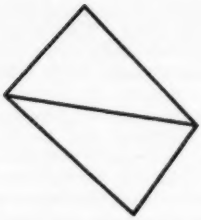


FIG. 2. The wire "parallelogram" used in Exp. 2

for a comparatively short period, namely as long as it forms a contour of the shadow, whereas in the realistic situation it usually can also be seen when it passes across the front of the figure. We therefore performed quantitative experiments with wire figures which of course show all edges continuously, and had them turn back and forth through an angle of only 42°.

Figures 2 and 3 each show projections of two of the wire figures used. The figure represented in Fig. 2 can best be described as a parallelogram containing one diagonal, which was bent along this diagonal so that the planes of the upper and the lower half formed an angle of 110° with each other. The other figure (Fig. 3A) consisted of a piece of wire twice bent to form part of what might be described as a triangular helix. Figure 3B shows a view from the top. These figures were turned back and forth through an angle of 42° at a rate of one cycle per 1.5 sec. and their deforming shadows were shown to individual *S*s one at a time for as many periods of 10 sec. as were necessary to obtain a clear report. After each 10-sec. exposure, *S* was asked for a report.

In the case of the parallelogram (Fig. 2) all of 50 *S*s sooner or later during the exposure periods reported seeing a three-dimensional figure turning back and forth apparently very much like the wire figure behind the screen, except that it sometimes appeared to *S* in its inverted form. In the case of the "helix" 48 of the same 50 *S*s reported a three-dimensional form like the wire figure and 2 *S*s saw a flat zigzag line which distorted before their eyes, a process which corresponds to the changing pattern of stimulation. Evidence that the three-dimensional forms which were perceived were due to the kinetic depth effect and that none of the projections represented in the deforming shadow by itself would have been seen as three-dimensional was obtained in the following way. To a new group of 22 *S*s five such projections of each figure were presented individually in random order. They were the projections of the two extreme positions at which the figures stopped and started to turn the other way and of three intervening positions chosen to be at about 10° of rotation apart. None of these stationary projections looked three-dimensional to any one of the 22 *S*s. Positions intermediate between the ones presented resembled them so closely that it seems impossible that they would lead to a radically different perceived form.

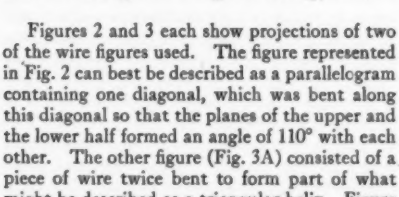


FIG. 3. The wire "helix" used in Exp. 2 (A) and its appearance from the top (B)

Often the turning wire figures were seen three-dimensionally immediately upon presentation. Of 16 *S*s for whom time records were taken 11 reported the "parallelogram" in the correct three-dimensional shape after the first 10-sec. exposure period. In the case of the "helix" only 4 of 16 *S*s reported its shape correctly after the first exposure. The reason for this difference is yet to be investigated.

Once it has occurred, the three-dimensional impression is so strong that one cannot voluntarily see the two-dimensional figure which is given

on the screen. It has been mentioned that the perceived three-dimensional figure may resemble the inverted form of the wire figure rather than that figure itself. This is to be expected, inasmuch as a three-dimensional figure and its inverted form have identical projections, such that a given projection is always a representation of both, a figure and its inverted form. The added fact that with prolonged observation of the distorting shadow the perceived three-dimensional figure may spontaneously invert in Necker cubelike fashion is also a consequence.

Experiment 3: Rotation of a truncated cylinder.—A systematic investigation of the kinetic depth effect resulting from rotation of a solid was attempted in the following manner.

A cylinder which is rotated about its axis casts a shadow which does not change in any way, but any cut that is taken off the cylinder and which is not at right angles to its axis produces a characteristic deformation of its shadow. A large number of wooden cylinders, about as high as wide, were made and cuts were varied systematically. The solid forms which were produced in this fashion were shown in complete rotation to *Ss* who had no previous experience with our studies.

The results can be summarized in the following way: (*a*) Shadows whose only deformation consists in an expansion and contraction in one dimension will look flat; a dark figure is seen which periodically becomes wider and narrower. An example is the shadow of a rectangular block which is rotated about an axis parallel to a set of edges. (*b*) Shadows which display contour lines that change their direction and their length will appear as turning solid forms. The roof-shaped figure described above (Exp. 1) is an example; the shadow contour produced by an edge of a gable tilts and changes its length simultaneously (compare Fig. 1B and 1C). (*c*) Curved contours which are deformed without displaying a form feature which identifies a specific point along the curve are seen as distorting, often even if for some reason the shadow is seen as a three-dimensional form. This peculiarity is in disagreement with our description of the kinetic depth effect and has delayed our work for years. It is now clear that the perceived distortions of deforming curved contours are not related to the kinetic depth effect at all. They will be dealt with in a later paper.

Experiment 4: Rotation of straight rods.—When it became clearly recognized that shadows which display contour lines that change their direction and length will appear as turning solid forms, we checked whether this would also apply to single lines. Will a detached line which changes its direction on the screen and its length at the same time display a kinetic depth effect, i.e., will it appear to tilt into depth in such a way as to account for its shortening? The following experiments show that this is the case.

When a rod is fastened to the end of a vertical shaft at an angle of, say, 45° and is rotated about this shaft, the shadow which it produces tips from side to side. From a tilt of 45° toward the left, the shadow rights itself and goes through the vertical into a tilt of 45° toward the right and back through the same positions. At the same time it shortens and lengthens periodically. Its end points move on horizontal lines across the screen. This process is invariably seen as a motion in three dimensions, as nearly as one can tell exactly like the real movement of the rod behind the screen, a rotation describing a conical surface about a vertical axis.

This set-up lends itself to a variation which is interesting because the three-dimensional motion which is perceived is a different one from that

which the rod behind the screen undergoes. The shaft on which the rod turns is tilted toward or away from the screen by 15° or 20°. This changes the pattern of expansion and contraction of the rod shadow completely; its end points now move on elliptic paths on the screen. Only rarely does the rod motion seem to describe a circular cone like the real movement of the rod behind the screen. In most cases a movement is seen in which the movement component of tilting toward or away from *S* is much smaller than the lateral movement component such that the surface described is that of an elliptic cone. Whereas the real rod goes in one rotation from a tilt toward *S* into a tilt away from him, the perceived motion of the rod is restricted either to a changing forward tilt or to a changing tilt away from *S*. Either one of these two different movements can of course be perceived because, like a three-dimensional figure and its inverted form, they would if they were objectively given produce the same projection.

It seems that what distinguishes in these experiments perceived movement from the two-dimensional process which is given on the screen and therefore on the retina is the fact that the perceived rod is seen with a constant length. Tilting into depth is seen instead of shortening. Inasmuch as the rod as a whole must anyway be seen to move, this tilting motion is only a modification of a necessary process and not an added change, which is what a perceived stretching and shrinking would amount to. Thus, a tendency to see one motion instead of the two simultaneous movements of the two-dimensional process (a tipping from side to side, and a stretching and shrinking) may be held responsible for the depth effect. Another possibility would be a selective principle according to which a line of constant length is seen rather than a changing one. No decision between these two possibilities can be made on the basis of our present results.

Experiment 5: Rotation of T *and* Δ *figures.*—It is important to realize that a shadow line must undergo both a displacement and a lengthening or shortening in order to produce a kinetic depth effect. Both these changes must be given together. A change in length alone is not sufficient to produce a reliable kinetic depth effect. This is shown in the following experiment.

A piece of 1/16-in. wire 4 in. long was fastened at right angles to a vertical shaft of the same diameter so that the two formed a figure of T shape. When this form was turned back and forth about the vertical shaft through an angle of 42°, the shadow of the horizontal wire formed a line of 105-mm. length which periodically contracted to a length of 75 mm. and expanded again. We presented it with a rate of one period per 1.5 sec. for 10 sec. to 24 *Ss*.

Eighteen of 24 *Ss* saw a line in the plane of the screen expanding and contracting and only 6 *Ss* perceived the line turning in the third dimension. Had the horizontal wire been in an oblique position with respect to the axis of rotation so that its shadow had shown a displacement in addition to the change in length, the kinetic depth effect would have been obtained with the majority of the *Ss*, as the experiment with the oblique rod would predict.

Because of its importance we had the experiment with the T figure repeated by another *E* and obtained the same result: Of 40 *Ss*, 30 saw the horizontal line expand and contract in the plane of the screen and 10 saw it turn.

These results must be compared with those of a different wire figure which does produce a kinetic depth effect under identical conditions. Comparable data come from a shadow presentation of an equilateral triangle. Its sides consisted of wires 5 in. in length, and one of them was tilted by 15° against the horizontal. When this figure was turned back and forth, the shadow of one of its sides changed

from a slope of 45° to one of 57° and thus presented conditions favorable to the kinetic depth effect. Of 20 *Ss* who observed the triangle for 10 sec. 17 reported a turning in depth and only 3 expansion and contraction. (With such a plane figure the kinetic depth effect consists, of course, in the perception of a plane figure that turns into depth.) This score is to be compared with that for the T figure where only one-third of the *Ss* saw a turning. The difference is reliable at the .001 level of confidence.

Still, the 6 and 10 *Ss* who saw the T figure turning in depth need be accounted for, if our claim is correct that they are not the outcome of a genuine kinetic depth effect. Here the following result is significant. When the T figure was presented *following* the presentation of the triangle, 15 out of the 17 *Ss* who saw the triangle turn saw the T figure turn also. (Three *Ss* who saw the triangle expand and contract reported the same for the T figure.) In other words, when a figure that *Ss* saw turning preceded the presentation of the T figure, a large majority of *Ss* saw it turning also. This shows a strong influence of a previous perception on the manner in which the T figure is seen, for only one-third of the *Ss* saw the T figure turning when it was given as a first presentation. This difference is significant at the .05 level of confidence. We suggest that in the case of these latter *Ss* some such influence of a previous experience, though a more remote one, has been at work also.

We are inclined to conclude that a line which changes in length but is not displaced at the same time does not give rise to the kinetic depth effect. This agrees well with the already reported finding that a solid shadow which expands and contracts only in one dimension will not show the kinetic depth effect either.

The Accuracy of Kinetic Depth Perception

When we described the kinetic depth effect in complex figures, we stated that the change in the retinal image is accounted for in perception by a rotation of a three-dimensional form. This implies, of course, that a *real* form which is like the perceived one would in rotation produce a sequence of retinal images very similar to that actually given, or, in other words, that the perceived form resembles closely the shadow-casting object. That this is the case has been confirmed by many *Ss* to whom the shadow-casting object was directly shown immediately following the kinetic presentation.

For more stringent confirmation several methods were used: In the case of the "helix" (Fig. 3), *Ss* were asked to bend a piece of wire into the shape of the figure which they saw turning on the screen. Only *Ss* who reported seeing a three-dimensional form turning back and forth were given this test. Of 29 *Ss*, 13 made good reproductions, 12 fair ones, and only 4 *Ss* made poor reproductions. Eleven of the 12 *Ss* whose reproductions were fair were later asked to make another reproduction while they looked *directly* at the turning wire figure. Of these only 7 were able to make good reproductions and 4 made only fair ones under these conditions. Altogether there were 8 *Ss* who could make only fair reproductions when they viewed the figure directly.

In the case of the parallelogram (Fig. 2), the accuracy of perception by virtue of the kinetic depth effect was checked by showing *S* four similar wire figures and asking him to pick out the one that matched best the form he saw turning on the screen. As mentioned earlier, the bend in the shadow-casting figure amounted to an angle of 110° between the planes of the two triangles which made up this figure. This angle is, of course, characteristic of its three-dimensional form. If this angle were 180°, the wire figure would be plane. In our four models the bend amounted to angles of 95°, 110°, 125°, and 140°, respectively. The one with the 110° angle was an exact copy of the shadow-casting

figure; the other three models had the same height as the standard and were made to produce projections on the frontal plane which were identical with the projection of the standard. The four models were inserted in a wooden block and handed to *S*. The choices of 30 *S*s who had previously reported seeing the three-dimensional figure were 8, 17, 4, and 1 for the 95°, 110°, 125°, and 140° figure, respectively. Only one *S* found it impossible to make a choice. It is unfortunate that we did not include one more model with a still sharper angle, but even so the data give a rough idea of the accuracy of the perception of three-dimensional form which is based on the kinetic depth effect.[3]

Experiment 6: Rotation of luminous rod.—Another attempt to obtain a measure of the accuracy with which depth perception functions by virtue of the kinetic depth effect was made employing a straight line which rotated in an oblique plane. When a line is turned in a frontal-parallel plane about its midpoint, the end points of its retinal projection move on a circle. However, when it rotates in an oblique plane, its retinal image changes in length as it turns, and its endpoints move on an elliptic path. Therefore, when in the dark a luminous line is rotated in a frontal-parallel plane about its midpoint and is exposed from behind an elliptic aperture, its retinal projection will be the same as that of a line which turns in an oblique plane, for the aperture causes the line to be visible with the same changes in length. If the motion of the line that rotates in an oblique plane can be correctly perceived with the help of the kinetic depth effect alone, then the line rotating behind an elliptic aperture should also appear to rotate in an oblique plane, because the two lines produce identical stimulation so far as the kinetic depth effect is concerned. Other cues for depth perception as, for instance, retinal disparity, would give rise to experienced rotation in an oblique plane only in the first case. Thus when the rotating line behind the aperture is presented to a naive *S* and he perceives it turning in a properly oblique plane, one can be sure that this is due to the kinetic depth effect. By this procedure, as by the use of the shadow screen, the effect can be studied in isolation; other cues for depth perception would tend only to prevent the line from turning in an oblique plane.

A ½-in. lucite rod, approximately 23 in. long, served as light source for the luminous line. Its ends were flat and finely ground to admit a maximum of light. They were inserted in metal caps which contained hidden flashlight bulbs and also served as mountings. The light from these bulbs made the lucite rod appear to glow evenly over its whole length. The rod was inserted in a U-shaped sheet-metal trough of proper length and attached by the mountings. A bushing was fastened to the back of the trough at its midpoint and attached to the horizontal slow shaft of a reduction gear motor, so that the lucite rod could be turned in a vertical plane. The trough and rod were covered by a long strip of cardboard into which was cut an aperture 22 in. long and ¼ in. wide. Through this aperture part of the rod's surface was visible. In front of this apparatus was a frame to which cardboards with different elliptic apertures could be attached parallel to the plane of rotation of the rod. Three different elliptic apertures were used. They were 40 cm. long and 35 cm. wide, 39.5 cm. long and 28.9 cm. wide, and 40 cm. long and 20.6 cm. wide, respectively. They produced projections of the turning luminous rod identical with the projections produced by a luminous line that turns in a plane forming an angle of 29°, 46°, and 59°, respectively, with the plane of the aperture. They were attached with the large axis of the elliptic opening in vertical position.

The *S* was seated in front of this apparatus at a distance of 9 ft. He had before him a small table covered with a light gray cardboard. A metal rod was joined at right angles to a shaft which was fastened vertically to the table, so that the rod could be swung around in a plane parallel to the table top about an inch above it. A degree scale was marked out on the cardboard by which the position of the rod could be read. With the help of the rod, *S* could indicate the position of the plane in which the luminous line

[3] It should be noted that a turning by 42° produces only a moderate distortion of the shadow. Its width which is most strongly affected suffers a reduction of only 27%.

seemed to turn. A darkroom amber bulb illuminated this arrangement in such a way that *S* could see the rod but not the scale markings and that the remainder of the room was completely dark. After each setting of the rod *E* asked *S* to close his eyes and then read the scale with a flashlight.

During testing *E* asked *S* to look at the luminous line before him and close one eye, and that eye was covered. The luminous line which, when presented in the appropriate elliptic aperture, produced the projection of a 46° tilt was set into clockwise rotation at a rate of one revolution in 10 sec. When *S* reported that it turned in an oblique plane, he was asked to turn the measuring rod before him into a position parallel with that of the luminous line in rotation. This was repeated with the 59° and the 29° apertures, and thereafter the three apertures were used twice more in random order for the purpose of practice. Neither in this practice period nor later during the experimental trials was *S* told whether his settings were correct or not correct. No time limit was set on the presentation of the revolving line and *S* made his setting when he felt ready. After a rest period of 3 or 4 min. the experimental series began. It consisted of nine presentations, that is, each one of the three apertures was presented three times in random order.

The means and *SD*'s of the 15 means of the three measurements for each of the three degrees of tilt were 14.9 (*SD* = 6.6), 44.0 (*SD* = 3.4), and 59.8 (*SD* = 4.4) for the 29°, 46°, and 59° tilt, respectively. There was *no* overlap between the means of the settings by individual *S*s from one tilt to another, and there was *no* overlap between the individual settings which a given *S* made for the three tilts.

Where projections of tilts of 46° and of 59° were presented, the averages for all *S*s came close to the expected values. However, in the case of the 29° tilt, the average of 14.9 deviates significantly from this value; only one individual setting out of 45 is as high as, or exceeds 29°.

Why this is so is not clear. However, it should be pointed out that the change in length which a line turning with a 29° tilt undergoes amounts only to 12.5% and one of 14.9° tilt (the value of the average) only to 3.4%. In other words, a tilt of 14.9° produces a change in length that is very likely below the limen. Yet, that does not necessarily mean that those *S*s who gave settings of 15° or lower did not receive effective stimulation for a tilt. Whereas the small angles of tilt (due to the negligible change of the cosine function in this range of values) probably do not lead to a change in length sufficient to produce a kinetic depth effect, settings of such low values do not indicate that no tilt was perceived when these settings were made. In experience, a tilt of 15° is of distinct significance, and objectively conditions of stimulation of 29° were given.

It would have been important to find out whether these results can be improved by making the luminous line wider. To make it wider would have the advantage that the contraction of the line would be given by a change in its proportions, that is, in figural terms, rather than by a change in its absolute length. Improved results would indicate that change in proportions is effective in producing the kinetic depth effect. Unfortunately, this variation could not be done with the present set-up, because a wider line would have shown up the intersections with the elliptic aperture by their changing obliqueness. A more expensive manner of presentation would be needed.

Other Factors in Kinetic Depth Perception

Experiment 7: The effect of angle constancy.—Such a variation of the experiment might have contributed to the solution of the following problem. It has been reported that shadows of solid blocks will produce the kinetic depth effect only if they have contours which are displaced and change their length simultaneously. Should we assume then that the presence of such contours solely accounts for the kinetic depth effect in complex forms, imparting their depth to the whole figure, or do complex forms produce such an effect in their own right? Just as we have considered the possibility that a line is seen to move into the third dimension to account for the given change of its retinal image while it is perceived with constant length, we might assume a tendency to see in

FIG. 4. The figure used in Exp. 7

general rigid, unchanging forms instead of the given distorting shapes. For the present this question must remain unanswered.

However, a question which can be considered a part of the question just raised was actually put to a test. Most shadows of turning figures do not only display contours which are displaced and change their length; their angles also change. Is there a separate tendency for angles to remain constant which produces kinetic depth effects?

To answer this question we used a figure which consisted of three rods all meeting in one point under angles of 110° and forming a wire-edged representation of an obtuse corner (Fig. 4). When this figure was turned back and forth through an angle of 42° and its shadow was shown to 56 *Ss* who before had seen its stationary shadow as two-dimensional, 53 *Ss* reported seeing a rigid obtuse corner. In this presentation two of the three dark lines forming the shadow were not only displaced but also underwent considerable changes in length.

Entirely different results were obtained when the length of the shadow lines was made indefinite. To achieve this, the shadow screen was covered with a cardboard with a circular aperture where the shadow of the corner figure fell on the screen. The size of the aperture was so chosen that the ends of the shadow lines were always hidden from *S*. Thus, only movement of the corner point, angular displacements of each of the three lines, and changes of the angles which they formed with each other were visible. As the corner point shifted sideways, one of the lines seemed to move farther under the aperture edge and another one seemed to pull out from under it, and the length of all three of them seemed indefinite.

All 22 *Ss* employed in this experiment reported seeing a flat figure which distorted. Had the kinetic depth effect occurred, the *Ss* would have seen instead a rigid three-dimensional form with constant angles. However, no such effect was observed where changes in the length of the lines which constituted the figure were not given, because that length was indefinite. We may conclude that a displacement of lines which is linked only with a change of angles does not give rise to the kinetic depth effect; displacement of lines and change in their length is needed. The question remains unanswered whether length must be understood only in absolute terms, or whether change in proportion has an effect of its own.

Experiment 8: Variation of distances between objects.—Not only are the retinal images of solid objects deformed when one moves about, the same is true to various degrees of the projection of the whole environment. That the objects which make up the environment are seen arranged in three-dimensional space and with unchanging distances between each other may also result from a kinetic depth effect. Just as some of the contours of solid objects produce appropriately changing retinal projections when the objects are seen from different angles, the projections of many of the intervals between objects change their length and their direction when one moves about. That this has the effect of producing visual depth can be demon-

strated with the shadows of an arrangement of several objects on a rotating platform.

We used spheres supported by thin vertical rods because the shadow of a sphere in rotation does not change its shape and will therefore not produce a kinetic depth effect of its own. Four spheres of 1 3/16-in. diameter were arranged at the corners of a square concentric with the platform. The four rods were all of different height so that all the intervals between the shadows of the spheres were periodically oblique when the platform turned, and changed length and direction simultaneously. The lower part of the screen was covered so that the shadow of the turntable itself was hidden. The arrangement was turned at a rate of one revolution in 5 sec. and was exposed to individual *S*s for 20 sec.

Under these circumstances all 30 *S*s who took part reported the spheres to move in three dimensions. Twenty-four *S*s saw the spheres in a rigid spatial arrangement which turned about its center just like the actual arrangement behind the screen. The others saw them move in open single file in snakelike fashion into depth. In this latter motion only the shorter intervals between spheres are rigidly maintained, and each sphere changes direction of rotation with each excursion; but there can be no doubt that this is an incomplete form of a kinetic depth effect. It should be mentioned that 15 of the 30 *S*s had no other experience with these shadow experiments except for another experiment with the spheres which will be reported below. It is apparent that the kinetic depth effect will readily yield a perception of a rigid spatial arrangement of unconnected objects.

Experiment 9: Variation of distances between objects.—The arrangement of Exp. 8 offers still another opportunity to check on our finding that a line must change both in length and in direction in order to produce a kinetic depth effect. From our experiments with solid blocks and with the T figure we concluded that a shadow which merely expands and contracts in one dimension will not give rise to the effect. Here we set out to show that the same is true for intervals between objects; we modified the experiment with the spheres to correspond to these experiments.

All the rods were given the same height so that the spheres were aligned on a horizontal line and the intervals between the shadows changed in length only. This arrangement was shown to the same 30 *S*s and under exactly the same conditions that prevailed in Exp. 8, but was presented prior to it. The fact that the T figure proved so susceptible to the influence of previous perceptions made this sequence advisable. The experiment with the luminous rod which has been reported at length was done with the same *S*s in between the two experiments with spheres.

As in the case of the T figure, only a minority of the *S*s now saw movement in three dimensions, namely 10 out of 30. The difference between this result and that of Exp. 8 is reliable at the .03 level of confidence. Of the 15 naive *S*s, 12 saw the shadows move back and forth in the plane of the screen, 1 *S* reported movement in three dimensions, and 2 saw in the beginning of the observation period the plane and later the three-dimensional version. For the other group of 15 *S*s who had observed the shadows of some wire figures before, the numbers were: 8 flat, 2 three-dimensional, and 5 first flat and later three-dimensional. When these results are compared with those of the previous experiment where imaginary lines connecting sphere shadows changed both in length and in direction, it becomes again apparent that these are essential conditions for the kinetic depth effect and that a mere expanding and contracting of retinal distances is insufficient.

This is the reason for our view that Metzger's work (2) is not directly con-

cerned with the kinetic depth effect. He exposed shadows of arrangements of vertical rods whose changing patterns presented *Ss* with rectangular intervals which changed in width only. No other deformations were visible, because no marks were distinguishable along the shadow lines, and the latter ended only at the edges of the aperture in which they were given. As with the aligned spheres, no reliable depth effects are produced spontaneously in naive *Ss* with such an arrangement. Whether Metzger's work contributes to an understanding of the kinetic depth effect remains to be seen when more of the nature of the effect is known.

Experiment 10: Effect of set.—In our experiments, the kinetic depth effect results in two perceptual characteristics: (*a*) a turning in the depth dimension and (*b*) three-dimensionality of form. However, when the effect is observed under realistic conditions, namely when by moving about one obtains a changing retinal projection for a stationary solid object, no turning is perceived. The reason for this is that the object remains in unaltered relation to its environment with respect to which *S* perceives himself moving. Thus, only three-dimensionality and rigidity of form, seen instead of the deforming two-dimensional pattern which is given on the retina, are here the overt manifestations of the kinetic depth effect.

The fact that under realistic conditions the object remains in unaltered relation to its environment needs some consideration, because this is not so in our experiments. There the deforming shadow of the object denotes a turning while the environment, that is, the screen on which the shadow is shown, remains stationary. Under realistic conditions, on the other hand, both the object and its environment are given retinally with deformations which denote a turning in relation to *S*. The kinetic depth effect transforms the deforming retinal projection of the environment into a three-dimensional structure, and once this has happened, the perception of the object as a three-dimensional form is probably facilitated. That such a facilitation is likely to take place is indicated by experimental results which show a strong influence of preceding exposures on the readiness with which the kinetic depth effect occurs.

The following results may serve as an example. When the "helix" (Fig. 3) was shown following the presentation of one or two figures which readily show the kinetic depth effect, all of 18 *Ss* saw it as three-dimensional during the first 10-sec. exposure. This is to be compared with results in Exp. 2 according to which only 4 out of 16 naive *Ss* gave a clear report of three-dimensional form for this figure after the first 10-sec. exposure.

The question of how this influence is exerted must remain open. It is conceivable that it consists merely in a set to see a turning in the depth dimension. However that may be, the influence is a strong one.

If such an influence is effective between succeeding exposures of different figures as shown, it should be expected to work also within a given visual field. Under realistic conditions, once the environment of a given object is perceived as a rigid spatial structure which changes its orientation with respect to the moving *S*, such an influence should facilitate a kinetic depth effect for the object. There are several reasons why the environment should easily be seen in this fashion. To mention only one: the environment will usually contain familiar features which would cause the facilitating influence of previous experience with similar situations to operate. Thus we have good reason to believe that the kinetic depth effect takes place more readily under realistic conditions than it does in our shadow-screen experiments.

Summary

When a three-dimensional form, solid or wire-edged, is turned behind a translucent screen and its shadow on the screen is observed, the shadow will appear as a rule as a three-dimensional rigid object which turns, quite similar to the physical object behind the screen. This happens notwithstanding the fact that *S* actually looks at a plane figure which is being deformed.

One condition seems to be essential for the occurrence of this effect: the shadow must display contours or lines which change their length and their direction simultaneously. If this condition is not fulfilled, a plane distorting figure like the one on the screen is perceived unless an influence of previous perception operates.

This effect is believed to operate widely under ordinary circumstances. When one moves about, objects near one's path are successively seen from different angles, and this change in orientation of the object to *S* is the same as occurs when the object is turned by an equivalent angle. Thus, the object's retinal projection undergoes the same deformations as do shadows in our experiments, and the same perceptual processes should result.

(Received August 11, 1952)

References

1. Fisichelli, V. R. Effect of rotational axis and dimensional variations on the reversals of apparent movement in Lissajous figures. *Amer. J. Psychol.*, 1946, **59**, 669–675.
2. Metzger, W. Tiefenerscheinungen in optischen Bewegungsfeldern. *Psychol. Forsch.*, 1934, **20**, 195–260.
3. Miles, W. R. Movement interpretation of the silhouette of a revolving fan. *Amer. J. Psychol.*, 1931, **43**, 392–405.
4. Philip, B. R., & Fisichelli, V. R. Effect of speed of rotation and complexity of pattern on the reversals of apparent movement in Lissajou figures. *Amer. J. Psychol.*, 1945, **58**, 530–539.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

SUSTAINED BEHAVIOR UNDER DELAYED REINFORCEMENT[1,2]

CHARLES B. FERSTER

Harvard University

In some preliminary research the rate of responding of pigeons pecking at a small illuminated disc was examined under delays of reinforcement of from 1 to 10 sec. It was found that the initial effect of a delay was a decrease in the rate of responding. When delay procedures were continued over longer periods, however, the frequency of pecking increased to a magnitude that might be obtained with no delay. The present experiments determine some of the conditions which influence the rate of a response whose reinforcement is delayed.

Experiment I

Problem.—A study of the effect of long delays in reinforcement when *S*s have had no special training.

Method.—The *S*s of the experiment were four adult pigeons. The experimental cage was a box which was lightproof and attenuated about 20 db. The front panel of the box contained a grain magazine and an illuminated 1-in. circular hole behind which was a hinged Plexiglas assembly, the key. Pecking the key closed a contact and permitted recording and programming of the experiment. Operation of the magazine presented an illuminated hopper of grain to the bird for 4 sec.

A variable-interval schedule whose mean was 1 min. was used throughout the experiment. The intervals in the series ranged from 3 to 120 sec. in an arithmetic progression. These intervals, arranged in a random series, produced a steady rate of response by which the effect of delays in reinforcement could be measured. The top part of Fig. 1 illustrates a variable-interval schedule of reinforcement. The arrows above the line define the intervals in variable-interval schedule, each arrow representing the "setting up" of a reinforcement. The first response after a reinforcement has been "set up" opens the magazine.

The box was darkened for the duration of the delay period in order to prevent *S* from responding. The bottom part of Fig. 1 illustrates the delay procedure. Reinforcements are available on the 1-min. variable-interval schedule mentioned previously. The reinforced response turns out all the lights in the box. With the return to normal illumination after the delay interval, the magazine opens. Neither the recorder nor the reinforcement programmer operates during the delay interval. Four *S*s were run on the variable-interval schedule until their rates of response were stable. On the same reinforcement schedule, reinforcements were delayed by 60 sec.

Results.—Figure 2A is the cumulative record for one *S*. The segment of the curve before the arrow records the behavior without delay in reinforce-

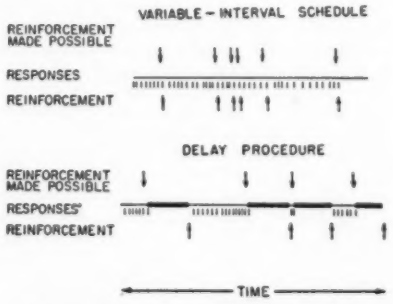


Fig. 1. The top part of the figure illustrates a variable-interval schedule of reinforcement. The bottom part of the figure illustrates the delay procedure. The darkened portions represent the delay interval.

[1] This experiment is part of a research program carried out under Contract N5ori-07631 between Harvard University and the Office of Naval Research, U. S. Navy (Project NR 143–943, report PPF-2), directed by B. F. Skinner.

[2] Thanks are due to B. F. Skinner for suggestions given during the course of the research. Thanks are also due to J. A. Dinsmoor, M. Kaplan, and J. G. Beebe-Center for reading and commenting on the manuscript.

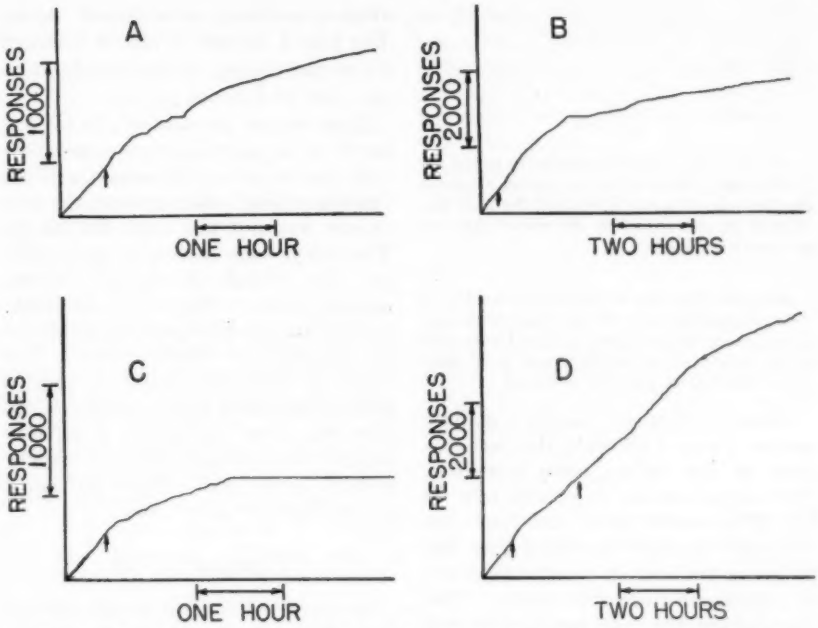


FIG. 2. Cumulative records of four pigeons showing the effect of a first encounter with a 60-sec. delay. The segment of the curve before the arrow records the stable state of a variable-interval schedule of reinforcement. At the arrow, reinforcements are delayed by 60 sec. At the second arrow, in D, the delay was increased to 120 sec.

ment; it is typical of all prior behavior. Beginning at the arrow, where all reinforcements were delayed by 60 sec., the rate of responding becomes progressively lower until the rate falls below one-sixth the original value. Figures 2B and 2C show a second and third *S* under the same procedure. In all three cases the rates of pecking decline substantially after the introduction of the delay in reinforcement. The record in Fig. 2D is of a bird that was able to maintain a nearly normal rate of response under a 60-sec. delay in reinforcement. The 60-sec. delay (beginning at the first arrow) causes a small decline in rate. With a delay of 120 sec., however (beginning at the second arrow in Fig. 2D), the rate falls off to about one-half the original value. The four *S*s were continued under the delay procedure for approximately 30 hr., during which the rate continued to decrease.

Since these birds were unable to maintain a rate of response appropriate to the schedule of reinforcement under 60-sec. delays, any subsequent ability of the birds to maintain high rates of responding under 60-sec. delays must depend specifically on the training procedures employed.

EXPERIMENT II

Purpose.—A study of the training by which pigeons were enabled to maintain high rates of response under long delays in reinforcement.

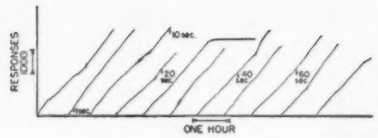


FIG. 3. The complete cumulative record for one *S* showing the result of a program of gradually increasing delays. The size of the delay was changed at the arrows to the values that are given on the record.

Method.—The four *S*s that were unable to maintain normal rates of responding with long delays under the conditions of Exp. I were subjected to a program which began with short delays which were gradually increased.

Results.—Figure 3 is the complete record of one *S* showing the performance as the delays were increased. The record covers approximately 90 hr. The curves show that there are no major changes in rate during the 90 hr. of training in which the delay was increased from 1 to 60 sec. The rate during the last segment of the curve where the delay was 60 sec. does not differ from the first segment where the delay was 1 sec.

One bird afforded a dramatic example of learning to bridge the delay. Figure 4 shows the effect of introducing a 60-sec. delay in reinforcement after a training on a 40-sec. delay. The initial decline in rate is followed by an exceedingly smooth acceleration to a rate of 3 pecks per sec.

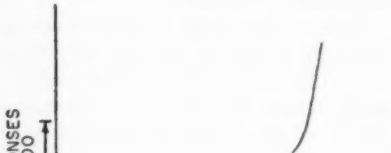


FIG. 4. A dramatic example of learning to bridge the delay: the effect of a 60-sec. delay in reinforcement after training on a 40-sec. delay

Since some amount of "art" was involved in judging when it was possible to increase the delay without "extinguishing" the response, the procedure was not the same for all *S*s. The delays were increased too rapidly for the fourth *S*. Under 60-sec. delayed reinforcement it did not maintain a rate of response appropriate to the schedule of reinforcement. The three *S*s that did maintain normal rates of response under 60-sec. delays were continued under the same procedure for several hundred experimental hours each. They exhibited no tendency to slow down.

EXPERIMENT III

Purpose.—A control experiment on the effect of the blackout without reference to delay in reinforcement.

Method.—The choice of a method by which the reinforcements were delayed was arbitrary. It would have been possible to slide a gate over the key during the delay period, for example. The blackout was chosen, however, and the question arose as to whether any of the extinction effects observed when *S*s were first exposed to long delays were due to the blackout without reference to its delaying function. Accordingly, a control experiment was carried out to find the effect of the blackout when it was not related to the reinforcement.

Three birds were conditioned on a 1-min. arithmetic variable-interval schedule of reinforcement until the rates of pecking were stable. Three different procedures were applied: (*a*) a variable-interval schedule with no blackout; (*b*) a variable-interval schedule with 60-sec. blackouts occurring independently of both responses and reinforcement; (*c*) a variable-interval schedule with blackouts occurring as a consequence of a response but independent of reinforcements. The occurrence of the blackouts was determined on a variable-interval schedule identical with the reinforcement schedule but out of phase with it. Each *S* underwent some 100 hr. of experimentation during which the three procedures were

rotated. Six hours were spent on a procedure before the next was introduced.

Results.—The resulting records did not differ from each other and it was concluded that the effect of the blackout (not correlated with the occurrence of reinforcements) is simply to introduce a period of "dead time" into the experiment.

Experiment IV

Purpose.—A study was made in order to obtain direct observation of the behavior which occurs during the delay period. This was designed to determine what factor in Exp. II accounts for the persistence of normal rates of pecking under 60-sec. delayed reinforcements. A reasonable suggestion is that some form of "superstitious" behavior becomes conditioned during the delay period, which bridges the gap introduced by the delay (**1, 2**).

Method.—To make it possible to check on the presence of a "superstitious" delay response the procedure of Exp. III was changed so that the box was illuminated continuously. When a reinforcement was available (on the same variable-interval schedule that was used in the preceding experiments), a peck turned out the key light for 60 sec. (but not the general box light). At the end of the 60-sec. period, the magazine opened and the light behind the key came on again. If *S* pecked during the delay period, the opening of the magazine was postponed an additional 60 sec. In this manner it was impossible for any peck to precede the opening of the magazine by a period greater or less than 60 sec. As is usual in the formation of a discrimination the frequency of pecking the darkened key soon dropped to zero.

Results.—Observation of *S*s through a one-way screen revealed that each *S* had acquired a characteristic delay response which it executed consistently and at a moderate rate during the delay period. In two *S*s the delay response was "turning in a circle with the head stretched high." From 35 to 45 turns were made during an 80-sec. delay period. The third *S*'s "superstition" was "pacing back and forth in front of the key panel with the head stretched high." The recording of such behavior offers a serious problem of instrumentation since the response which will be acquired during the delay period cannot be predicted. In this experiment it was sufficient to observe the birds visually. The "superstitious" delay responses occurred consistently over an experimental period of approximately 100 hr., during which the pecking response was maintained at a rate appropriate to the schedule of reinforcement.

Experiment V

Purpose.—The purpose of this study was the demonstration of a chain analogous to the delayed-reinforcement chain. Another method of checking the preceding analysis of delayed reinforcement is to demonstrate that a chain which is deliberately conditioned by *E* can function in the same way as the "superstitious" chain conditioned as a consequence of the delay procedure. This was done by setting up a chain in which two members could be recorded.

Method.—The apparatus was similar to the above except that either a red or a blue lamp could be illuminated behind the key. A discrimination was established between the red and the blue light. The first response that occurred 1 min. after the blue light came on was reinforced (S^D) and the key color changed to red. No responses were reinforced in the red (S^Δ). After a variable interval in red, without reference to *S*'s behavior, the color changed to blue and the cycle repeated. After 4 hr. the rate of pecking the red key fell to zero, while the performance on the blue key was characteristic of a fixed-interval schedule of reinforcement.

The procedure was now changed so that the key color did not change from S^Δ to S^D unless the bird responded during S^Δ. After 1 hr. of continuous reinforcement of the S^Δ responses by a change to S^D, the schedule of reinforcement in

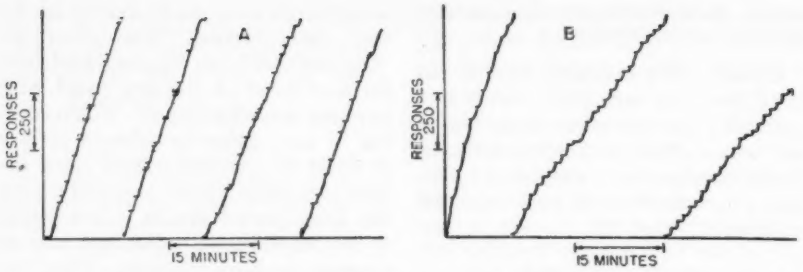


FIG. 5. Cumulative records for one *S* for the two members of the chain. The two recorders operated alternately; the small horizontal lines mark reinforcements. (*A*) *Red key:* Responses are reinforced on a 1-min. variable-interval schedule by a change to blue. (*B*) *Blue key:* Responses are reinforced on a 1-min. fixed interval by the operation of the magazine.

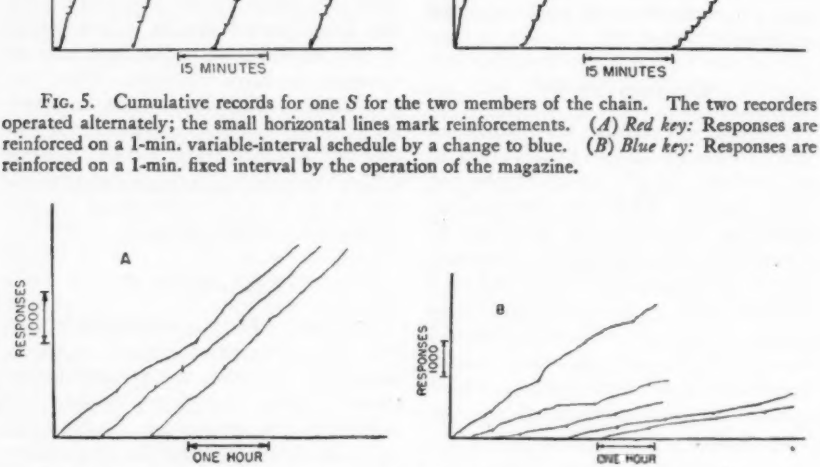


FIG. 6. Cumulative records for the two members of the chain for one *S* where the magazine opened whether or not *S* pecked the blue key. (*A*) *Red key:* Responses are reinforced on a 1-min. variable-interval schedule by a change to blue. (*B*) *Blue key:* The magazine opens 1 min. following the change to blue whether or not *S* pecks.

S^{Δ} was changed to a variable-interval schedule. A chain of responses was now required as follows: Pecking the red key was reinforced on a variable-interval schedule by a change to blue. On a 1-min. fixed-interval schedule responses to the blue key were effective in opening the magazine. This situation differs from the discrimination procedure only in that the occurrence of S^{D} is contingent on a response in S^{Δ}. The two members of the chain were recorded separately with tandem recorders. When one recorder operated, the other did not. The difference between the S^{Δ} rate under the discrimination procedure and the S^{Δ} rate under the chain procedure must be ascribed to the reinforcing effect of the discriminative stimulus.

Results.—Figures 5A and 5B show the records for the S^{Δ} and S^{D} behavior, respectively, for one pigeon. The behavior in the presence of each key color is appropriate to the schedule of its reinforcement. The second member of the chain which is on a fixed-interval schedule of reinforcement shows the typical pause after reinforcement and the acceleration to the final rate. The first member of the chain which is on a variable-interval schedule of reinforcement shows a constant rate. Thus each part of the chain derives its character from the schedule of reinforcement appropriate to that member via the discriminative control of the key color. To make the chaining procedure analogous to the delayed-reinforcement procedure, it was necessary only to break the contingency between the opening of the magazine and the response to the blue

key. Pecking the red key still changed the key color to blue on a 1-min. variable-interval schedule as before. The magazine opened 1 min. after the key turned blue whether the bird pecked or not. The red key responses which produced the S^D (blue) correspond to the pecking responses in the delayed-reinforcement experiment. The blue key responses—the magazine opens independently of the response—correspond to the "superstitious" behavior that took place during the delay period in the delayed-reinforcement procedure. Figures 6 and 7 show the S^Δ and S^D record, respectively, for the chaining procedure in which the magazine opens without reference to whether *S* pecked the blue key. The S^Δ records are typical of the performance on the regular chaining procedure and remain stable. The S^D rates, however, show a smooth decline to a lower but stable value even though the response reinforcement contingency is superstitious.

Systematic Analysis

The preceding experiments have shown that a delay in reinforcement is not a static parameter of the effect of reinforcement on behavior. The effect of a delay on the frequency of a response depends critically on the way in which the bird is introduced to the particular delay. By introducing *S* to short delays which were gradually increased to 60 sec., it was possible to maintain normal rates of response under long delayed reinforcement. If the first delay had been of this magnitude, the response would have extinguished.

Reasons for the maintenance of normal response rates in spite of the long intervening period between the response and the appropriate reinforcing stimulus may be found first, in an examination of the behavior taking place during the delay period, and second, in the chaining of the delay responses (1).

Any behavior that occurs at the end of the delay period will be reinforced by the opening of the magazine, under the discriminative control of the darkness. The relation between this behavior and the opening of the magazine is spurious, since the magazine will open whether or not the response is made. This kind of behavior has therefore been called "superstitious." The "superstitious" behavior is not likely to be weakened by its occurring when the light is on because the pecking response is prepotent by virtue of its longer history of reinforcement. Any extinction of the "superstitious" response will, therefore, occur in the dark. Where the first delay in reinforcement was of the order of 60 sec., there was ample time during the delay period for the extinction of the response which had been conditioned in the previous delay period.

Because of the extinction taking place during the 60-sec. delay the opening of the magazine is preceded by a wide variety of behaviors and no single response is likely to be reinforced frequently enough to acquire any strength. If the delays are of the order of 1 to 5 sec., however, the likelihood of the same response occurring prior to the opening of the magazine is high, and members of a single response class will be reinforced frequently enough to be conditioned. When the delay interval is lengthened after a number of reinforcements of the same response, the delay defines a fixed-interval schedule of reinforcement of the superstitious response. The schedule is different from those ordinarily employed only in that the magazine will open whether or not the response is made. To maintain the "superstitious" behavior in the dark, it is necessary to adjust the interval of the delay so that the magazine will open when the frequency of the "superstitious" response is highest.

The properties of a fixed-interval schedule of reinforcement will have a crucial bearing on whether a rate of occurrence of the "superstitious" response will be high at the time that the magazine opens. A fixed-interval schedule produces different effects after long training, when a stable performance is reached, than it does early in training. In the early stages of fixed-interval reinforcement the frequency of the behavior is high at the beginning of the interval and declines regularly until the next reinforcement occurs. These conditions are adverse for the maintenance of the "superstitious" behavior unless the delay is adjusted so that the magazine opens at a time when the rate of response due to the preceding reinforcement is still high. After sufficient training, however, the rate picture is reversed. The rate is low at the beginning of the interval and

gradually increases until a stable rate (characteristic of the fixed interval) is reached. Under these conditions the delay interval can be lengthened more rapidly without producing many instances of the opening of the magazine being preceded by a response other than the one previously conditioned.

If some response becomes conditioned under the discriminative control of the delay, stimulus pecking will be immediately reinforced by the beginning of the delay period. The strength of the pecking response will not be maintained directly by the presentation of food, but depends on the food-reinforced mediating behavior. A chain of responses will have been established as follows: Pecking the lighted key is reinforced on a variable-interval schedule by darkness, which is the occasion upon which some superstitious response is effective in opening the magazine. The schedule of reinforcement of the second ("superstitious") member is necessarily a fixed interval whose length is equal to the delay of reinforcement. It makes no difference that the dark response does not open the magazine in a strict sense. So long as the behavior is followed by the opening of the magazine sufficiently often, the resulting contingency will be sufficient to maintain the behavior in the same way that the usual fixed-interval schedule does.

Summary

The effect of delayed reinforcements was shown by using a variable-interval schedule of reinforcement as a baseline. It was found that in three out of four pigeons the effect of a 60-sec. delay in reinforcement was a decline in the rate of pecking to a value near zero. The fourth *S* maintained near normal rates of responding under 60-sec. delayed reinforcement, but slowed down under 120-sec. delays.

The same *S*s that extinguished under 60- and 120-sec. delays were reconditioned on the original variable-interval schedule without delay. Short delays were then introduced and the value of the delays was gradually increased as *S*'s behavior became stable under any particular delay. It was possible in three out of four of the cases to sustain normal rates of response under delays of 60 sec.

A control experiment showed that the effect of a delay in reinforcement is not due to the aversive properties of the blackout by which the reinforcements were delayed.

Visual observation of *S*s under the delay procedure showed the presence of "superstitious" behavior during the delay interval.

The function of the "superstitious" behavior in the delayed reinforcement was investigated further by setting up a chain in which two members could be recorded which could function in the same way as a "superstitious" chain conditioned as a consequence of the delayed procedure.

The factors governing the effect of a particular delay and the effect of the change from one delay to another are the same as those determining the formation of a chain of responses.

(Received August 18, 1952)

References

1. Keller, F. S., & Schoenfeld, W. N. *Principles of psychology.* New York: Appleton-Century, 1950.
2. Skinner, B. F. Superstition in the pigeon. *J. exp. Psychol.*, 1948, **38**, 168–172.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

THE INFLUENCE OF EXPERIMENTALLY INDUCED FAILURE ON THE RETENTION OF MATERIAL ACQUIRED THROUGH SET AND INCIDENTAL LEARNING

MURRAY ABORN

Michigan State College

In his review of the literature on experimental studies of repression, Zeller (6), concludes that no adequate laboratory test of repression has been made and proposes an experimental design to test the hypothesis that "repression is a process which inhibits recall without destroying memory for the subject matter." Briefly, Zeller's method involves the following steps: (*a*) Equated experimental and control groups learn some experimental material to the same criterion. (*b*) Following a time interval, both groups are measured for retention. No difference in performance should obtain. (*c*) Experimental *Ss* are subjected to some "ego threat" associated with the learned material; control *Ss* are subjected to some neutral procedure. (*d*) Both groups are measured for retention. Control *Ss* should perform significantly better. (*e*) The previously induced "ego threat" is alleviated for the experimental *Ss*. (*f*) Both groups are measured for retention. Experimental *Ss* should demonstrate a "restoration of memory" (no particular criterion for "restoration" is specified).

In two later experiments, Zeller (7, 8) concludes that an experimental analogue of repression was achieved by the application of his method, although he notes some qualifications are in order in terms of variables other than repression, e.g., overlearning and motivational factors. The present study attempts to apply Zeller's basic methodology, with certain modifications in design and technique, to the investigation of the effects of experimentally induced failure on retention for material learned by *Ss* with a set to learn and identical material learned by *Ss* with no set to learn, i.e., incidentally learned material. Although previous experimental studies of repression have recognized the existence of learning set as a variable, specific investigations of the influence of learning set have largely been confined to "set" in the sense of a previously existing attitude or "frame of reference," as Edwards (3) has proposed. Set in this sense includes factors lying outside of *E*'s control. Consequently, the failure to retain is likely to involve selective perception or differential forgetting rather than repression.

The present study considers set to learn in the nature of Zeller's experiments, where *Ss* were required to learn material in which their own preconceptions were not necessarily involved. A comparison will then be made between the retention curve of such set-to-learn *Ss* and the retention curve of *Ss* acquiring the experimental material incidentally. Changes in the retention curves under the influence of the introduction and alleviation of an "ego threat" will constitute the measure of repression.

METHOD

Materials.—A total of 19 color plates were selected from the Ishihara (4) and Dvorine (2) color vision tests. Of these, nine were taken from the former and ten from the latter. Of the nine Ishihara plates, four show numerals, two reveal a numeral only to the color-blind, and

three show no numeral whatsoever. All the Dvorine plates show numerals, making a total of 14 plates with numerals and five nonnumeral plates. The plates with numerals were chosen to represent figures as scattered and dissimilar as possible. Each plate was mounted on a separate sheet of cardboard for ease of manipulation.

Six designs were selected from the Arthur Stencil Design Test I (1).[1] The designs chosen were of a low or moderate level of difficulty and four of them required Stencil No. 13 for reproduction. By removing Stencil No. 13 from the testing materials, four of the six test problems became impossible to solve.

The memory tests consisted of a blank sheet of lined paper for recording numerals recalled and a paper-and-pencil recognition test in which the 14 color-plate numerals appear randomly among a list of 32 numbers. Of the 18 irrelevant numbers in the list, many are reversals of the 14 pertinent numerals or numbers sequentially proximate to them.

Procedure.—Eighty *S*s were randomly assigned to four groups of 20 *S*s each: Set to Learn—Experimental, Set to Learn—Control, Incidental Learning—Experimental, and Incidental Learning—Control. Each *S* went through the testing procedure individually. A Set to Learn—Experimental *S* entered a small room equipped only with a table, chairs, and a desk lamp. The *S* was told by *E* he was going to take a color vision test. The nature of the color test was briefly described and *S* was instructed to call out and then to try to memorize each numeral he saw. The *S* was told he would later be asked to recall them. The color plates were then presented in a prearranged order, *S* being permitted to examine each plate for 5 sec. The 14 plates with numerals were followed by the five nonnumeral plates, *E* meanwhile keeping a record of *S*'s responses. If *S* inquired about the absence of numbers on the nonnumeral plates, he was told to report a number if he saw one, or none if he saw none. When *S* had gone through the set of plates once, he was told he would be given two more trials, each under a different illumination. He was again instructed to call out and to try to retain each number he saw. The overhead light was turned off and the desk lamp turned on; dim for the second trial, bright for the third.

When the three color-plate trials were concluded, *E* recited a statement informing *S* that the purpose of the experiment was to prove that a relationship exists between color vision and "certain aspects of intellectual performance." The statement was designed to imply: (*a*) that people with good color vision have high ability in "certain intellectual areas" while those with poor color vision are of low ability in these areas, and (*b*) that the results of the color test indicate that *S* is "color-weak." The *S* was told he would next be given a test to examine his ability in intellectual areas related to color vision.

The *S* was escorted to another room where another *E* was prepared to administer the six Arthur Stencil Designs. Stencil No. 13 had been removed from the testing materials beforehand. *S* was given the test instructions and permitted 1 min. on each item. On those items that were now nonsolvable, *E* stopped *S* at the end of the time limit and directed him to proceed to the next item.

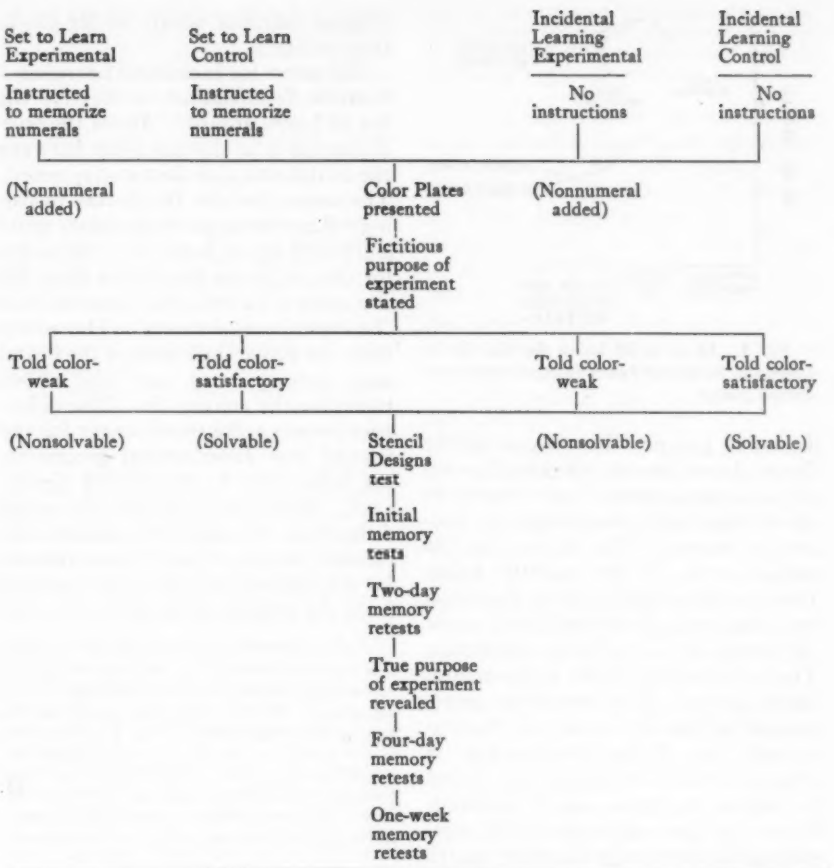
At the conclusion of the Stencil Design test, *S* was given no statement about his performance (his failures being quite apparent) but taken to a third room where another *E* administered the memory tests. The *S* was asked first to write down every color-plate numeral he could recall. When that was done, the recall sheet was removed from sight and *S* was directed to check off every numeral in the recognition list he definitely identified as belonging to one of the color plates.

The *S*s were contacted unexpectedly two days later and given both memory tests again. At the end of the retest, *S*s were informed of the true purpose of the experiment. They were told of the ruse of the nonnumeral color plates and the nonsolvable Stencil Designs. The *S*s were encouraged to voice their reactions to the experiment, most of them signifying that they had taken the results seriously. They were again assured that their color vision was normal and that the Stencil Designs were deliberately made incapable of solution.

Two days later, *S*s were again contacted and retested with the memory tests. They were again retested unexpectedly at the end of another three days.

The procedure for Set to Learn—Control *S*s was identical to that for the experimental *S*s except for the following: (*a*) *S*s were not shown the five nonnumeral color plates. To account for the difference in testing time, they were delayed in the testing room for a period equivalent to the time consumed by the experimental *S*s in viewing the nonnumeral plates during the three trials. (*b*) After being given the fictitious statement of experimental purpose, *S*s were told their color vision was "satisfactory." (*c*) Stencil No. 13 was included in the testing materials so that all the items were solvable. The *S*s were again delayed in the testing room to compensate for any difference in time as compared with that

[1] This subtest of the Arthur Point Scale of Performance Tests, Revised Form II, consists of 20 3-in. square symmetrical color designs and 18 paper stencils of different color and pattern. Each design can be reproduced by placing some combination of stencils, one atop the other, in the correct order.



consumed by the experimental *S*s (who worked to the time limit on the nonsolvable items).

The procedure for the Incidental Learning—Experimental *S*s was identical to that for the Set to Learn—Experimental *S*s except for the complete omission of any instructions regarding the retention of the color-plate numerals at the beginning of the experiment. The procedure for the Incidental Learning—Control *S*s was identical to that for the Set to Learn—Control *S*s except for the complete omission of any instructions regarding the retention of the color-plate numerals at the beginning of the experiment.

The experimental procedure is summarized in the above paradigm.

Subjects.—All *S*s were undergraduate students drawn from several elementary-level psychology courses. There was a slightly greater number of males than females among the *S*s, but about the same proportion appeared in each of the four groups. Data obtained from any *S* found truly color-blind during the course of the experiment would necessarily be excluded from the experimental results. Fortunately, all the *S*s participating in the study turned out to be adequate in this respect.

Results and Discussion

The experimental procedure and results are depicted graphically in Fig. 1 and 2. Figure 1 shows the general trends of the retention curves as described by the mean recall scores for the Set to Learn and Incidental

FIG. 1. Mean recall scores for the Set to Learn and Incidental Learning experimental and control groups

Learning groups. In the case of the Set to Learn groups, the recall scores of both experimental and control *Ss* show progressive decrements on successive retests. The scores for the experimental *Ss* fell slightly below those for the control *Ss* after the initial test, but none of the individual retest differences are statistically significant. The introduction of the experimental threat seems to have been of no consequence for the Set to Learn—Experimental *Ss*. Their performance is almost identical to that of the control *Ss*, where no threat was introduced. It may be that experimental *Ss*, possessing the knowledge that they would be required to recall the numerals and placed under threat by *E*'s fictitious statement, immediately inferred that their ability to recall was to be taken as an index of their intelligence. In that event, it is more than likely that these *Ss* attempted to improve their memorization during the remaining experimental procedures by rehearsal. The failure of any evidence of repression to appear may thus be due to a tendency to compensate for the effects of repression. This may be regarded as a general criticism of all unsuccessful repression experiments where the original learning occurs under conditions of set.

The curve for Incidental Learning—Control *Ss* is similar to that for the Set to Learn groups. About the only difference is a sharper drop between the initial test and the two-day retest. The curve for the Incidental Learning—Experimental *Ss* manifests quite a different trend, however. The scores for this group are lower than those for the control *Ss* until the alleviation of the experimental threat. The scores then rise above their level at the initial and two-day tests, and rise above those for the control *Ss*. The difference between the recall scores for the control and experimental groups on the initial test is statistically significant. While none of the remaining differences are similarly statistically reliable, there is at least some evidence for the existence of repression, judging from the graphic difference in trends.

Figure 2 shows the general trends of the retention curves as described by the mean recognition scores for the Set to Learn and Incidental Learning groups. In the main, the trends are in keeping with those shown in Fig. 1. The slopes of the curves for the Set to Learn groups are somewhat steeper than they were on the recall tests, but the differences and the trends are the same. The generally lower magnitude of almost all the recognition scores is due more to a correc-

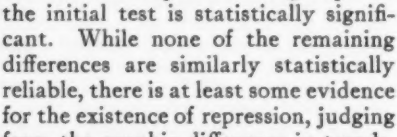


FIG. 2. Mean recognition scores for the Set to Learn and Incidental Learning experimental and control groups

tion-for-guessing formula employed in scoring the recognition tests than to any difference in memory, per se. Some correction for guessing was deemed necessary in view of the limited number of choices provided by the test, and if anything, the comparatively lower corrected recognition scores indicate a tendency, on the part of the Incidental Learning—Control *Ss*, to have attempted more guessing than did the Experimental *Ss*. The significant difference obtained for the Incidental Learning groups in the initial recall test does not occur among the recognition scores. No significant results appear for individual test differences among the retention retests, but this was also true for the recall scores. Graphically, the same difference in trend among experimental and control groups appears.

To this point, the results suggest that experimentally induced failure acts differently upon retention when the material learned was acquired with the absence of instructions to memorize than when it was acquired under conditions of set. In the case of the latter, the induced failure seems to have induced no concomitant repression in the form of either initially disrupted performance or later recovery. Both experimental and control *Ss* demonstrate simple forgetting with time. In the case of *Ss* with no instructions to memorize, the control group behaves in a similar way. The experimental group, however, evidences not simple forgetting, but initial disruption of performance, a tendency to retain what was recalled after failure was induced, and a tendency toward greater recall once the threat of failures was alleviated. Whether this trend toward "restoration" of memory constitutes an index of repression is ultimately a matter of how much recovery is to be expected. Experimental *Ss* cannot be expected to regain all of what might have been remembered had no threat been introduced since control *Ss*, representing the magnitude of recall without threat, demonstrate progressive forgetting with time. Some loss, then, is to be anticipated from experimental as well as control *Ss*. The most reasonable criterion would seem to be threefold: that experimental *Ss* score significantly lower than control *Ss* following the introduction of threat; that experimental *Ss* equal or surpass control *Ss* once the threat is alleviated; that following the alleviation of threat, experimental *Ss* demonstrate a significant trend toward recovery *based upon their own level of performance at the time threat was introduced*. It might be pointed out that the third condition of the criterion is unnecessary since the recovery of memory can only be viewed from the relative performance of control *Ss*, and since this is implicit in the second condition of the criterion. It is still felt, however, that an air-tight test of repression, in its classical meaning, would involve some overt evidence of increased recall, particularly where experimental and control groups were not prematched.

In keeping with the above criterion, three statistical hypotheses remain to be tested: (*a*) for each group independently, a test of the significance of the score changes among the several retention tests; (*b*) for each group independently, a test of the significance of the linear trend; (*c*) provided the first two procedures yield significant results, a test of the regression of each linear trend upon the other, particularly where the trends are in opposition.

The results of the independent trend analyses are presented in Table 1.[2] Two *F* tests are reported in the case of each group: the significance of the variance attributable to score changes among the retention tests when tested against the interaction variance, and that of the linear trend among the retention tests when tested against the pooled error. The linear effect represents the sums of squares of the differences among the four retention tests, adjusted for the unequivalence of the retest time intervals.

The differences and linear trends on successive retests are significant for the Set to Learn *Ss*. It will be recalled that the experimental and control groups performed at about the same

[2] The analysis of variance given in Table 1 is based upon smaller *N*'s than were included in the experiment and in the results so far. Whereas 80 *Ss* went through the experimental procedure and there were 20 *Ss* in each group on the initial test, some *Ss* later proved unavailable for all of the retests. Consequently, the *N* for each retest fell off according to the number of *Ss* that were obtainable. By the time all four tests were completed, the number of *Ss* having taken all of them was as follows: 13 Set to Learn—Experimentals, 9 Set to Learn—Controls, 8 Incidental Learning—Experimentals, and 9 Incidental Learning—Controls. Only these *Ss* were included in the above analysis.

TABLE 1

ANALYSIS OF VARIANCE

Group	Source of Variation	Set to Learn Groups					Incidental Learning Groups				
		df	Recall		Recognition		*df*	Recall		Recognition	
			Sum of Squares	*F*	Sum of Squares	*F*		Sum of Squares	*F*	Sum of Squares	*F*
Experimental	Total	51	368.00		600.98		31	200.00		149.89	
	Tests	3	26.15	8.63**	52.83	9.42**	3	.75	—	2.74	—
	Individuals	12	305.50		480.73		7	169.00		99.46	
	T × I	36	36.35		67.42		21	30.25		47.69	
	Linear effect	1	23.55	22.86**	49.92	26.98**	1	.04	—	1.46	—
	Other effects	2	2.60		2.91		2	.71		1.28	
	Error (T × I × Other)	38	38.95		70.33		23	30.96		48.97	
Control	Total	35	331.60		648.80		35	192.97		122.81	
	Tests	3	14.00	3.80*	47.60	4.81**	3	55.64	9.87**	5.86	—
	Individuals	8	288.10		522.50		8	92.22		66.43	
	T × I	24	29.50		78.70		24	45.11		50.52	
	Linear effect	1	13.97	12.25**	42.16	13.01**	1	42.58	19.01**	5.61	2.88
	Other effects	2	.03		5.44		2	13.06		.25	
	Error (T × I × Other)	26	29.53		84.14		26	58.17		50.77	

* Significant between .05 and .01 level of confidence.
** Significant at the .01 level of confidence or better.

levels and in the same downward trend in the case of both recall and recognition. No further analysis is necessary for these groups since it is clear that the introduction and alleviation of threat produced no significant differences between them.

The differences and linear trends on successive retests are significant for the Incidental Learning *Ss* in the case of the control group measured by recall only. The analysis reveals no significant upward trend on the part of the experimental group. Assessed in terms of the existence of repression for experimental *Ss* then, the criterion has been achieved on two counts. They manifest disruption in performance by scoring significantly lower than control *Ss* on the initial test, and they surpass control *Ss* on tests given after the alleviation of threat. They do not, however, show a significant improvement over their own initial level of performance, at least insofar as the length of time during which retests were made permits generalization. The results for the Incidental Learning—Experimental group are in some concordance with those reported by Russell (5), although any extensive comparison is precluded by the fact that his *Ss* underwent procedures of relearning and were initially set to learn. Without the opportunity for relearning the stimulus material, the Incidental Learning—Experimental *Ss* of the present study showed a strong tendency to retain the degree of memory manifested immediately following induced failure and a slight tendency to regain memory during later retests.

SUMMARY AND CONCLUSIONS

An attempt was made to determine the influence of experimentally induced failure, with consequent "ego threat," on retention for material

acquired with set to learn and identical material learned incidentally. The *Ss* were randomly divided into four groups: those given instructions to memorize the stimulus material without subsequent threat induced (Set to Learn—Control); those given instructions to memorize the stimulus material with subsequent threat induced (Set to Learn—Experimental); those given no instructions to memorize the stimulus material without subsequent threat induced (Incidental Learning—Control); and those given no instructions to memorize the stimulus material with subsequent threat induced (Incidental Learning—Experimental). All *Ss* were measured for retention by recall and by recognition immediately following the experimental procedure and after 2, 4, and 7 days. The induced failure was alleviated at the end of the two-day retest. Analysis of the retention curves yielded results leading to the following conclusions:

1. The introduction and alleviation of the threat produced marked differences between the performances of Set to Learn and Incidental Learning *Ss*. The threat had no measurable inhibitory effect on Set to Learn—Experimental *Ss*. In the case of Incidental Learning—Experimental *Ss*, the introduction of the threat produced significantly depressed memory scores and the alleviation of the threat resulted in some recovery. Whether the degree of recovery constitutes a valid index of repression depends upon what criteria one is willing to accept. Relative to the progressive decrease in retention exhibited on successive retests by the Incidental Learning—Control group, the experimental group showed no further decrease in retention beyond the initial retention test and ultimately surpassed the level of performance of the control group. The slight gain in retention observable for the experimental group following the alleviation of the threat was not significantly different from the level of their performance following the introduction of the threat, however.

2. Incidental-learning conditions would seem to have some advantages over set-to-learn conditions in the experimental study of repression. Without the knowledge that recall will be required, *Ss* are less apt to counteract any inhibitory effects of threat by rehearsal, overlearning, or increased motivation.

3. The differential results obtained for *Ss* placed under experimentally induced threat associated with material acquired with and without set to learn suggests an explanation of why certain aspects of an unpleasant experience are recalled with facility while other aspects remain largely inaccessible.

(Received August 18, 1952)

References

1. Arthur, G. *The Stencil Design Test I.* New York: Psychological Corp., 1947.
2. Dvorine, I. *Color perception testing charts.* Baltimore: Waverly Press, 1944. Vol. I.
3. Edwards, A. L. The retention of affective experiences—a criticism and a restatement of the problem. *Psychol. Rev.*, 1942, **49**, 43–53.
4. Ishihara, S. *Tests for color blindness.* Tokyo: Kanehara, 1925.
5. Russell, W. A. Retention of verbal material as a function of motivating instructions and experimentally-induced failure. *J. exp. Psychol.*, 1952, **43**, 207–216.
6. Zeller, A. F. An experimental analogue of repression: I. Historical summary. *Psychol. Bull.*, 1950, **47**, 39–51.
7. Zeller, A. F. An experimental analogue of repression: II. The effect of individual failure and success on memory measured by relearning. *J. exp. Psychol.*, 1950, **40**, 411–422.
8. Zeller, A. F. An experimental analogue of repression: III. The effect of induced failure and success on memory measured by recall. *J. exp. Psychol.*, 1951, **42**, 32–38.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

REMINISCENCE EFFECTS IN PAIRED-ASSOCIATE LEARNING

DONALD A. RILEY

University of California

One interpretation of reminiscence which has been frequently advanced and which has been able to account for much of the data on reminiscence assumes that the phenomenon is due to the differential rates with which competing responses are forgotten; weaker wrong responses presumably being forgotten at a more rapid rate than the correct response. Although the work of Ward (8), Hovland (2), and others (6) may be interpreted as supporting this hypothesis, to date there have been no studies in which the task of *S* has been to learn both correct and incorrect responses to the same stimuli for differing numbers of trials so that at the time of the rest interval competing responses of different strengths would be insured. Hovland (2) has suggested that the failure to obtain reminiscence in paired-associates learning may be because the technique of different pair orders on each trial breaks up the formation of remote associations which operate as competing responses. If this is a proper interpretation, then the introduction of competing responses into the paired-associates situation, as suggested above, should result in a reminiscence effect. It is the purpose of the following experiments to test this suggestion.

Three experiments were conducted. Experiment 1 investigated the effect of a rest interval on postrest learning when the correct responses were presumably stronger than the incorrect responses; Exp. 2 investigated the effect of the rest when the wrong responses were not learned; and Exp. 3 investigated the effect of the rest when the wrong responses were presumably stronger than the right responses. According to the above hypothesis, we should find a beneficial effect of the rest in Exp. 1, no effect or a slight negative effect in Exp. 2, and a negative effect in Exp. 3 which would be larger than that in Exp. 2. In the interests of clarity, the report of the first two experiments will be given first, followed by the report of the third experiment.

Experiments 1 and 2

Method

In the discussion to follow, E1 and C1 refer to the experimental and control conditions of Exp. 1 in which matched *S*s in the two conditions learned competing correct and incorrect responses before the introduction of the rest interval in E1. E2 and C2 refer to the corresponding conditions in Exp. 2.

Materials.—Pretraining materials were eight pairs of three-letter words selected so as to be maximally dissimilar to each other, and three boards filled with color squares. The experimental materials were two lists of eight pairs of nonsense syllables hereafter referred to as Lists A and B, the stimulus items being the same in each list, the response items changing from A to B. The syllables were taken from Glaze's (1) lists, the stimulus items and each set of response items averaging 15% association value. In the construction of the lists, no letters in common were allowed between a stimulus item and its two response items nor between the two response items. In addition, the item similarity within the stimulus list and the item similarity within both response lists was held to a minimum, as was the similarity between the stimulus list and the two response lists, and the similarity between the two sets of response lists.

Procedure.—In the pretraining period *S*s were given instructions concerning paired-associates learning and then given eight trials on a memory drum in which the stimulus term first appeared

TABLE 1

MEAN TRIALS TO SUCCESSIVE SYLLABLE CRITERIA ON LIST B AND COMPARISON OF DIFFERENCES BETWEEN DIFFERENCES TO SUCCESSIVE SYLLABLE CRITERIA IN EXP. 1 AND 2

Syllables Correct	Experiment 1					Experiment 2					Differences Exp. 1 and 2			
	E1	C1	σ_{diff}	t	p*	E2	C2	σ_{diff}	t	p**	Exp. 1–2	σ_{diff}	t	p***
1	6.90	6.60	.36	.86	.20	4.80	4.30	.26	1.92	.10	.20	.44	.46	.35
2	9.85	12.00	1.17	1.84	.05	8.40	7.45	.80	1.19	.30	3.10	1.42	2.19	.02
3	13.75	16.25	1.87	1.34	.10	12.05	10.10	1.32	1.48	.20	4.45	2.29	1.95	.03
4	18.20	21.10	2.08	1.39	.10	16.55	13.95	1.77	1.47	.20	5.50	2.73	2.01	.03
5	23.50	27.50	2.83	1.41	.10	20.80	19.35	2.01	.72	.50	5.45	3.47	1.57	.10

* p values in Exp. 1 are for one-tailed t test; $df = 19$.
** p values in Exp. 2 are for two-tailed t test; $df = 19$.
*** p values of Exp. 1 differences minus Exp. 2 differences are for one-tailed t test; $df = 38$.

alone for 2 sec., followed by the stimulus and the response terms together for 2 sec. The next stimulus term followed immediately. The time between trials was 8 sec. except for a 2-min. interval after the fourth trial, during which *Ss* named colors from the color boards at a 2-sec. rate. The lists appeared in six different orders.

Immediately after the pretraining period the experiment proper began. In Exp. 1 the *Ss* were first given four trials on List A, followed *without a break* by six trials on List B. After the sixth trial of List B, Group E1 had a 2-min. rest interval during which time they named colors, and then continued the learning of List B to a criterion of one perfect recitation.[1] Group C1 followed the same procedure as Group E1 except that after the sixth trial of List B the usual 8-sec. interval occurred instead of the 2-min. color-naming period used in E1. In both groups, four different orders of List A and six different orders of List B were used to avoid serial effects. In addition to the instructions on how to learn the lists and the pretraining on the words, *Ss* in Exp. 1 were told that in the course of the learning some of the response terms would change and that they would have to make the best of the situation and learn whatever was present at any given time.

Conditions of Exp. 2 were exactly the same as Exp. 1 except that the four trials of List A were omitted, and the special instructions concerning the possibility of the response terms changing were not given.

Subjects.—The *Ss* in Exp. 1 were 40 elementary psychology students in the University of California Summer School. The *Ss* in Exp. 2 were 40 elementary psychology students in the University of California regular session. In each experiment *Ss* in the E and C conditions were closely matched on the basis of performance on the six trials of List B which preceded the rest interval in the E condition. Matching was sufficiently close that no pair of *Ss* differed by more than one syllable on any of the six trials, and no differences between pairs occurred on the sixth trial.

Results

Because all *Ss* did not learn to the criterion of one perfect trial in the 1 hr. allotted to the experiment but all did learn to a criterion of five out of eight syllables, the calculations are carried only to that point. Table 1 shows the mean number of trials on List B required by each group to reach the successive syllable criteria up to 5/8 correct. Although Groups E and C in both experiments were matched on the first six trials of List B, at which time the rest interval occurred in the E groups, the number of trials to each successive criterion (except for the first) favors E1 in Exp. 1, but favors C2 in Exp. 2. In each experiment the difference score between the

[1] The question may be asked why *Ss* were not required to learn to specified criteria (say, two out of eight on List A and three out of eight on List B) rather than for four and six trials as was done. The former undoubtedly would have been a superior procedure for several reasons. However, the latter was adopted in order that Lists A and B could be on the same tape, allowing the shift from A to B to occur with no time lag or warning activity on the part of *E*, thereby attempting to insure maximum interference between Lists A and B.

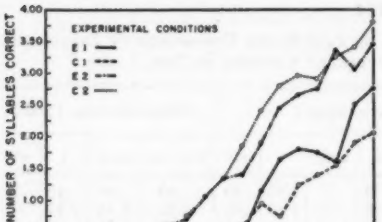


FIG. 1. Mean syllables correct on prerest and first eight postrest trials of Exp. 1 and 2

matched *Ss* was calculated at each successive syllable criterion, the means of these and their respective probabilities being shown in Table 1. In neither experiment are the successive means independent of each other, the general result of the lack of independence being to carry over the significance or lack of significance of a difference from one criterion to the later ones. Thus, in Exp. 1, the results at one syllable correct depress the significance of the results at two correct. If, instead of taking the number of trials to two syllables correct, we take the number of trials from one syllable correct to two syllables correct, the mean difference between conditions E and C increases to 2.45 trials, the σ_{diff} decreases to 1.10, and the t of 2.22 is significant at the .02 level of confidence. On the right-hand side of Table 1 is shown for each criterion the difference between the mean difference in Exp. 1 and the mean difference in Exp. 2 and the probability of each difference. Although these data suffer from the lack of independence mentioned above, they present satisfactory evidence of a difference between the differences in the two experiments.

Figure 1 shows the mean performance of the four conditions on the trials before the rest interval in the E conditions and the corresponding performance after the rest interval for all trials on which all *Ss* had scores. On the eighth trial after the rest interval one *S* recited the entire list perfectly, so the analysis was discontinued at this point. The difference between these two sets of differences was evaluated by a one-tailed t test, in which the variance due to trials had been eliminated from the denominator, and this difference was found to be of borderline significance ($t = 1.55$; $df = 38$). As indicated in Fig. 1, the differences are all in the predicted direction with the important exception that no reminiscence appears on the first postrest trial in Exp. 1.

Because of the possible significance of error commission for theories of distribution of practice (7), the errors in the two experiments were classified in three different ways and tabulated. The first classification, which applied only to Exp. 1, was of interlist intrusions from List A to List B. On the first trial of List B, there were 18 such errors in E1 and 19 in C1. They declined at equal rates and were almost completely gone in both groups by the sixth trial. There was no evidence of a recurrence of these errors after the rest interval in E1. Intralist errors in List B were tabulated in both experiments and were subjected to tests of significance because they showed the same trends as the error counts reported by Underwood and Goad (7). In Exp. 1 there were more errors in E1 and in Exp. 2 there were more errors in C2. However, as this difference did not remotely approach statistical significance ($F = .838$; $df = 1, 38$), the results have not been tabled. The third group of errors were those which could not be readily classified in other ways. The differences between Cond. E and C in the two

experiments were small and did not show appreciable trends in either direction.

It is apparent from the results in Table 1 that *Ss* in Exp. 1 learned more rapidly than those in Exp. 2. Data indicating that this difference is probably a result of the different source of *Ss* in the two experiments will be given in the analysis of the Exp. 3 results.

Discussion

In the introduction we stated that according to the differential-forgetting hypothesis if reminiscence is to occur, the incorrect responses should be weaker than the correct responses. While there is no guarantee that this requirement has been met, the procedure used and the absence of List A responses by the sixth trial of List B in Exp. 1 make it a reasonable assumption that the requirement has been met. From this it follows that the beneficial effect of the rest in Exp. 1 and the absence of such an effect in Exp. 2 confirm the hypothesis under test.

One datum, however, is not consistent with the differential-forgetting hypothesis, namely, the failure to find reminiscence on the first postrest trial in Exp. 1. The most likely explanation of this failure lies in the experimental naïveté of the *Ss*. The only other rest interval that *Ss* had experienced was the one during the pretraining period, and it is reasonable that the disruption might have masked any beneficial effect. Another possibility is that this postrest drop is characteristic of this general procedure. Hovland (2), working with experienced *Ss*, found similar decrements after a rest at three different levels of learning in the paired-associates task. The postrest results of Exp. 2 are in complete agreement with his, despite the differences in the method of determining the point of introduction of the rest interval and the slower learning of *Ss* in the present study. Experiment 1 differs from these experiments in the marked recovery of the rest group on subsequent trials. It is because of this difference that the first interpretation appears more plausible.

Within the last few years, several investigators (3, 4, 5, 7, 9, 10) have failed to find evidence supporting the differential-forgetting hypothesis, and as a consequence, at least two of the writers have suggested a performance-decrement hypothesis to account for reminiscence effects in verbal learning. McClelland (4, 5) has postulated inhibition which occurs as a function of competing reactions and the difficulty of the task. Hovland and Kurtz (3) have hypothesized a performance decrement which increases as a function of the mental work done in a given unit of time. In both cases, the work decrement is assumed to dissipate with a rest interval.

Applied to the present experiments, the work-decrement hypothesis would assert that a work decrement sufficient to reveal itself through beneficial effects of rest occurred in Exp. 1, where competing responses were learned, but that in Exp. 2 the lack of competing responses lessened mental work so that no reminiscence effects occurred. That the differences in mental work between the two experiments are due to the competing responses in Exp. 1 and are not due to the four additional prerest trials is suggested by Hovland's (2) test for reminiscence in the paired-associates situation. His *Ss* averaged 9.55 trials to reach a criterion of six out of nine syllables correct, the point at which the rest was introduced in the experimental condition. Although the number of prerest trials in the Hovland experiment is comparable to the number of prerest trials in Exp. 1, he found no beneficial effect resulting from rest.

Because this work-decrement hypothesis is a reasonable alternative to the differential-forgetting hypothesis in the explanation of the results of the two above experiments, a third experiment was conducted.

Again, on the basis of the differential-forgetting hypothesis the prediction was made that if the incorrect responses are stronger than the correct responses at the time of the rest interval, then the effect of the rest interval should be detrimental to the further learning of the correct responses. It is reasonable to assume that the effects of the rest in Exp. 3 should deviate from the effects of the rest in Exp. 2 to about the same magnitude as Exp. 1, but in the opposite direction.

Experiment 3

Method

The materials and the general procedures were exactly the same as in Exp. 1, the main difference being that the first four trials were of List B instead of List A, these being followed by six trials of List A instead of List B. At this point, the 2-min. rest interval was introduced in the experimental group (E3) and the usual 8-sec. rest followed the sixth trial of List A in the control group (C3). Following this, learning was continued on List B as in the earlier experiments. The *S*s were 40 elementary psychology students directly comparable to those in Exp. 2. They were matched on the basis of the six prerest trials of List A, and this matching was comparable to the matching on the earlier experiments.

Fig. 2. Mean difference between control and experimental groups to successive syllable criteria in Exp. 1, 2, and 3

Results

A comparison of the results of Exp. 3 with the results of the previous experiments is shown in Fig. 2 where the mean values of C minus E for the number of trials to each successive syllable criterion are shown for each of the three experiments. It is apparent that the results of Exp. 3 are in close accord with those of Exp. 1, and directly opposite the results predicted by the differential-forgetting hypothesis. Because the hypothesis under test would have justified the use of one-tailed *t* tests, had the differences between the mean differences been in the predicted direction, the use of two-tailed *t* tests to evaluate discrepancies in the opposite direction is not legitimate. Nevertheless, the probabilities from the two-tailed *t* tests have been recorded in Table 2 because of the striking similarity of the results in this experiment to the results reported in the earlier part of the paper.

Analysis of the number of syllables correct on each trial was carried out to the seventh postrest trial, the point at which one *S* completed learning. The results again duplicate those of Exp. 1, E3 becoming superior to C3 on the second postrest trial and remaining so for the rest of the trials. One important difference between the

TABLE 2

Mean Trials to Successive Syllable Criteria on List B in Exp. 3 and Comparison of Differences between Differences to Successive Syllable Criteria in Exp. 2 and 3

Syllables Correct	Experiment 3					Differences Exp. 2 and 3			
	C3	E3	σ_{diff}	*t*	*p**	Exp. 3–2	σ_{diff}	*t*	*p***
1	7.65	6.85	.60	1.33	.20	1.30	.65	2.00	.05
2	10.80	8.85	1.11	1.76	.10	2.90	1.65	1.76	.10
3	13.95	11.30	1.44	1.84	.10	4.60	1.95	2.36	.05
4	19.00	14.05	1.93	2.56	.02	7.55	2.61	2.89	.01
5	23.30	18.40	1.92	2.55	.02	6.35	2.78	2.28	.05

* *p* values in Exp. 3 are for two-tailed *t* test; $df = 19$.
** *p* values in Exp. 3 differences minus Exp. 2 differences are for two-tailed *t* test; $df = 38$.

results of the two experiments was noted. Speed of learning List B corresponds closely to learning speed in Exp. 2 rather than Exp. 1. The mean trials to 5/8 correct is 18.40 in E3 and 23.30 in C3. As these values bracket the corresponding values for Exp. 2, and as the pool of *S*s was the same for Exp. 2 and 3, the earlier assertion that *S*s for Exp. 1 were drawn from a different pool with respect to learning speed would seem to be justified. Had the differences between Exp. 1 and Exp. 2 in speed of learning List B been due to proactive inhibition in Exp. 1, then speed of learning List B should have been even slower in Exp. 3 because of retroactive inhibition. Since it seems quite clear that the difference between differences in Exp. 2 and Exp. 3 is due to the difference in the experimental operations, the previous assumption that the difference between Exp. 1 and Exp. 2 is also due to the experimental operations and not to the difference in learning speeds in the two groups is strengthened.

As in the previous experiments, errors were classified into interlist intrusions, intralist intrusions, and others. As in the other experiments, no apparent differences appeared between the interlist intrusions in the two conditions. There were a few more intralist errors in E3 than in C3, but the difference was less marked than in Exp. 1, where it did not approach significance. The unclassified errors in Exp. 3 showed a more marked frequency in E3 than in C3, and roughly corresponded to the intralist error differences in Exp. 1.

Discussion

If all three experiments are now considered together, it seems clear that the results are at odds with the differential-forgetting hypothesis, but may be interpreted as favoring a work-decrement hypothesis. To claim support for the differential-forgetting hypothesis, it would be necessary to argue that in Exp. 3 the original learning (List B) was stronger than the interpolated learning (List A) at the time of the rest interval. Values in Table 2 for prerest Trials 4 and 10 show that more syllables were learned in the interpolated learning, and thus this possibility is not supported.

Interpretation of the data in terms of the work-decrement hypothesis would seem to be quite straightforward. Learning one set of responses and then learning another set of responses to the same stimuli creates sufficient work decrement to produce a reminiscence effect. It is not clear from the evidence whether the critical factor in producing the work decrement is the performance of two difficult tasks within a brief period, whether it is the degree of interference between the tasks, or some other factor. The Hovland and Kurtz study (3) in which *S*s showed most reminiscence when switched without rest from multiplication problems to serial learning indicates the possibility that the learning of the two tasks may be sufficient to produce reminiscence, and the fact that the responses in Exp. 1 and 3 happened to interfere with each other is purely incidental. If this is the case, reminiscence would be expected even though the stimuli and the responses both changed, instead of just the responses as in the present experiments. On the other hand, if this work-decrement hypothesis is to account for reminiscence in serial learning, the beneficial effects of distribution of practice, and other related phenomena such as McClelland's (4, 5) finding of reminiscence in serial verbal discrimination learning, and Underwood and Goad's (7) recent finding that errors are positively related to beneficial effects of distribution in serial verbal learning, then the assumption that as interference increases, work decrement also increases would seem to be necessary. There is, of course, no reason why both a difficult task added to the main task and interference might not operate to increase work decrement.

Appealing as this work-decrement hypothesis might be, it is difficult to see how it is going to handle certain data such as the finding of the superiority of retention when learning has occurred under distributed practice and the results of the Hovland and Kurtz experiment (3) that the prerest learning was just as efficient in the work-decrement group as in the other groups. If release from work decrement improves performance, the presence of work decrement should be expected to harm performance. Although further investigations may bring discordant results like these into line with the theory, they do not seem to fit at the present time.

Summary

Three experiments were conducted to test the hypothesis that reminiscence is caused by differential rates of forgetting of competing correct and incorrect responses, the assumption being that reminiscence occurs only if the correct response is stronger, and thus more slowly forgotten than the incorrect response. All three experiments were paired-associates learning tasks. In Exp. 1, *Ss* first learned wrong responses for four trials and correct responses to the same stimuli for the next six trials. Following the sixth trial of the correct responses, a 2-min. rest occurred in the experimental group, the usual 8-sec. rest occurring in the control group. Learning of the correct responses then continued. In Exp. 2, *Ss* learned only the correct responses, a rest interval occurring after the sixth trial in the experimental group. After the rest, learning continued as in Exp. 1. In Exp. 3, *Ss* learned the correct responses for the first four trials, incorrect responses to the same stimuli for the next six trials, and then after the appropriate rests in the experimental and control groups, the learning of the correct responses continued as in the other experiments.

In Exp. 1, the group which received the rest interval learned more rapidly after the rest interval. In Exp. 2, the group which received the rest interval learned slightly slower than its control group after the rest interval. In Exp. 3, the group which received the rest interval learned more rapidly after the rest interval. The results of the first two experiments were in accord with the hypothesis under test. The results of the third experiment were directly contrary to the predicted results, the rest interval being beneficial as in Exp. 1. In both Exp. 1 and Exp. 3, noticeable reminiscence effects do not occur until the second postrest trial. Possible interpretations of the results are offered.

(Received August 23, 1952)

References

1. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
2. Hovland, C. I. Experimental studies in rote-learning theory: IV. Comparison of reminiscence in serial and paired-associate learning. *J. exp. Psychol.*, 1939, **24**, 466–484.
3. Hovland, C. I., & Kurtz, K. H. Experimental studies in rote-learning theory: IX. Influence of work-decrement factors on verbal learning. *J. exp. Psychol.*, 1951, **42**, 265–272.
4. McClelland, D. C. Studies in serial verbal discrimination learning. I. Reminiscence with two speeds of pair presentation. *J. exp. Psychol.*, 1942, **31**, 44–56.
5. McClelland, D. C. Studies in serial verbal discrimination learning. III. The influence of difficulty on reminiscence in responses to right and wrong words. *J. exp. Psychol.*, 1943, **32**, 235–246.
6. McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* (2nd Ed.) New York: Longmans, Green, 1952.
7. Underwood, B. J., & Goad, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–134.
8. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 4 (Whole No. 220).
9. Wilson, J. T. Remote associations as a function of the length of the interval between learning and recall. *J. exp. Psychol.*, 1943, **33**, 40–48.
10. Wilson, J. T. The formation and retention of remote associations in rote learning. *J. exp. Psychol.*, 1949, **39**, 830–838.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

A TEST OF THE MECHANISMS OF LEARNING PROPOSED BY HULL AND GUTHRIE

DAVID ZEAMAN AND LOUIS RADNER

University of Connecticut

Guthrie and Hull have different theories about how learning takes place. Guthrie (**2**) says that the necessary and sufficient condition for learning is the temporal contiguity (or overlap) of stimulus and response. Hull (**3**) would agree on the necessity of S–R temporal contiguity but he would insist that drive reduction or some stimulus surrogate of drive reduction be present to complete the sufficient conditions for learning. We have tried to devise an experimental arrangement in which these two views make different predictions.

The elements of the arrangement are simple. A rat in a box is presented with a stimulus situation including a bright light. He makes some arbitrarily selected response such as walking across the box, following which *E* turns off the light. Under these conditions the rat learns to cross the box when the light is presented. This much is already known (**5, 9**) and can be accounted for by either Guthrie or Hull. Hull would say that the bright light creates a drive state for the rat as well as illuminating the box stimuli. The crossing response is contiguous with the illuminated box stimuli and is followed by the drive reduction of the light-off. Thus there are present the necessary and sufficient conditions for learning a crossing response in the presence of light. Guthrie would hold that the experimentally forced correlation of light and crossing is a sufficient condition for the learning of the connection, the light-off serving merely as a big enough change in the stimulus situation to protect the light and associated stimuli from being conditioned to a new, non-crossing response.

Different predictions are made by the Guthrie and Hull theories when we reverse the conditions of the experiment. Now the rat starts from a dark box, runs, and his running turns *on* the light. Under these conditions we have the contiguity of dark-box stimuli and running, but no obvious drive reduction. According to Guthrie's principle we should expect the learned association of this S–R sequence, according to Hull we should not. It is our purpose to check the actual outcome by experiment.

Apparatus

A plywood box, 28 in. long, 8 in. wide, and 7½ in. high with a transparent lid made up the experimental chamber. It was painted white throughout. In order to better define a crossing response, the box was mounted on a center fulcrum so that a lengthwise crossing of the box would cause it to tilt through 4° of rotation. The tilt of the box also served to trip microswitches controlling a light source and timing clock.

The apparatus was located in a dark room. Illumination came from a 100-w. lamp whose light was directed into the box by a large mirror suspended over the box at a 45° angle. This arrangement prevented the box from being heated by the lamp and made easy the observation of *S*. Light intensity in the box was 31 ml. when the light was on, .001 ml. when it was off.

The apparatus was equipped with contacts, switches, and clocks such that either *E* could control the onset, duration, and termination of the light and dark periods, or control could be given to the rat in accordance with a procedure to be described.

The basic idea of this tilt-box light-aversion apparatus was originated by Keller and Oberlin (**5**).

Procedure

The two major experimental sequences are (*a*) light-to-dark (L-to-D), and (*b*) dark-to-light (D-to-L).

L-to-D.—Animals run under the first condition are manually placed in one end of the box with the bright light on. The light remains on until *S* crosses the box, whereupon the light goes off and stays off for 1 min. During the 1 min. of dark, *S* is free to do as he likes within the box. This completes one trial. The second trial begins immediately with the light going on regardless of which side of the box *S* is on. As soon as *S* crosses the box, the light goes off again for 1 min. to complete the second trial. Ten such trials are given each day. If learning takes place, the crossing response should become progressively more prompt when the light is presented. The measure to be taken is latency, i.e., how long it takes *S* to turn *off* the light.

D-to-L.—Rats run under the dark-to-light condition are also subjected to the above procedure but with conditions of light and dark interchanged. The *S* is started in the dark. As soon as he crosses, the light goes on and remains on for 1 min. Trial 2 then starts immediately with the termination of the light. Latency here is how long it takes *S* to turn *on* the light.

Extra responses.—Under Cond. L-to-D *S* is free to make crossings during the 1 min. of darkness in each trial. A record of these is kept. They are called extra responses. Extra responses in the D-to-L condition are made during the light period of each trial.

Forced responses.—After the start of a trial, *S* may fail to respond for a long time and hold up the experiment indefinitely. To prevent this we adopted the procedure of waiting for a given period of time (criterion latency—3 min. of no response, or in some cases, 1 min.) and then very gently nudging the unreactive rat with a finger until he made a crossing response. These were recorded as forced responses with criterion latencies.

General conditions.—Animals were run 23 hr. hungry at approximately the same time each night. A feeding of wet mash followed each nightly session by ½ hr. When *S*s were not in the experimental box, they were caged in the general colony room and there exposed for the most part to normal night-day light cycles.

Experimental design.—Each *S* served as its own control by running under both of the main experimental conditions. Eleven *S*s (Group I) ran for five days under L-to-D and then changed to D-to-L for five days. Five *S*s of Group I stayed under the D-to-L condition for a total of

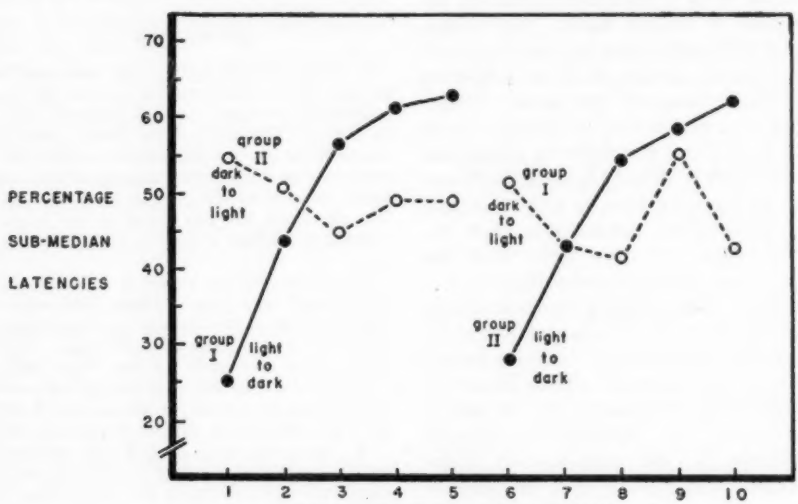


Fig. 1. The percentages of submedian latencies for Groups I and II under both conditions of light and dark sequence are presented as a function of the daily block of ten trials. The light-to-dark condition produces learning curves, the dark-to-light does not. There are 11 rats in Group I and 14 in Group II.

11 daily sessions. Group II (14 *Ss*) started under D-to-L, switching after five days to the L-to-D condition for five further sessions.

Subjects.—The *Ss* were a fairly heterogeneous group of 25 male albino rats, 17 of them at approximately 100 days of age, 8 at about 400 days. Eighteen of them were of Wistar strain, and 7 were Sprague-Dawley rats.

Results

The main findings appear in Fig. 1. The percentages of submedian latencies are plotted for both groups on each day. A rising function here denotes a progressive improvement or shortening of latencies as would be expected if the rats were learning to cross the box. For both groups the condition of running from light to dark yields regular learning curves. Contrasted with them are the curves representing the performance of rats running from the dark to light. These undergo no very regular course of change. Even after the six additional days of D-to-L training the subgroup of five rats from Group I failed to improve their performance. In fact, their latencies during Days 11 through 16 (not plotted) showed some slight tendency to become lengthened. The same effect can perhaps be seen in the slightly falling L-to-D curve of Group II. More evidence of a general weakness of responding under the D-to-L condition is available when we consider the matter of forced responses. First, however, the computations and rationale behind Fig. 1 should be given.

The recorded measures of performance were latencies. From these it seemed advisable to derive a frequency measure which would wash out two sources of variation in absolute level of latency not of immediate concern to our study. One source was from the wide individual rat differences in over-all level of latencies due probably to the wide range in ages, the other was from the activity-level differences known to exist between the conditions of running in the light and running in the dark (4, 9). A measure which compared each rat's latency on a given trial with his own median latency for the condition he was running under would be insensitive to these sources of variation but at the same time be sensitive to the latency changes produced by some process such as learning. In detail the measure was derived as follows. A median latency was found for each rat from the distribution of 50 latencies recorded for him during the five days under a single condition. A rat received a frequency score each day equal to the number of latencies *below* his condition median. The number of daily submedian latencies for all *Ss* was counted and divided by 10*N* (ten trials times the number of rats in the group) to give the plotted measure of *percentage of submedian responses.*

The data on forced trials show *Ss* running into the dark to be more responsive than those running into the light, although in general there were only a relatively small number of forced trials under either condition. The median number of forced trials per daily block of ten trials for each *S* was zero under both conditions. The mean number of forced trials under the L-to-D condition was .01; under D-to-L, 2.02. Other than this obvious difference between the two conditions there were no important trends in the forced-trial data.

The data on extra responses did not prove to be of significance for our major theoretical purpose. They are not reported.

Discussion

Taken at their face value the data support Hullian drive-reduction principles but disconfirm Guthrie's notion of the sufficiency of S-R contiguity for learning. The experimentally forced contiguity of the stimuli from a darkened box and a crossing response failed to produce evidence of learning, while the experimentally forced contiguity of the stimuli from a lighted box and a crossing response did produce and sustain learning when followed by the reduction of a light drive.

Before any general conclusion is accepted, two related questions must be answered: (*a*) Is this a fair test of Guthrie's theory? (*b*) Can an alternative explanation of these results be found without abandoning his contiguity principles?

With regard to the first question, it must be shown (*a*) that the conditions of the experiment conform to the variables specified in the theory, and (*b*) that the theory provides for the deduction of a definite behavioral outcome concerning which observations may be made. Guthrie has never formally organized his system, but Voeks (**8**) has formulated a set of definitions and postulates based on Guthrie's publications which are reported to be completely acceptable to Guthrie (**7**, p. 30). The postulates are important and brief, so they are reproduced in full:

Postulate 1: Principle of Association.—(*a*) Any stimulus-pattern which once accompanies a response, and/or immediately precedes it (by ½ sec. or less), becomes a full-strength direct cue for that response. (*b*) This is the only way in which stimulus patterns not now cues for a particular response can become direct cues for that response.

Postulate 2: Principle of Postremity.—(*a*) A stimulus which has accompanied or immediately preceded two or more incompatible responses is a conditioned stimulus for only the last response made while that stimulus was present. (*b*) This is the only way in which a stimulus now a cue for a particular response can cease being a cue for that response.

Postulate 3: Principle of Response Probability.—The probability of any particular response's occurring at some specified time is an increasing monotonic function of the proportion of the stimuli present which are at that time cues for that response.

Postulate 4: Principle of Dynamic Situations.—The stimulus pattern of a situation is not static but from time to time is modified, due to such changes as result from the subject's making a response, accumulation of fatigue products, visceral changes and other internal processes of the subject, introduction of controlled or uncontrolled variations in the stimuli present (**8**).

Applying these postulates systematically to the present experiment we have the following:

Light-to-dark condition.—During the light many responses can occur but the last one is always locomotion across the box. A combination of Postulates 1*a* and 2*a* predict the association of the lighted-box stimuli with crossing. Postulate 4 says there will be from trial-to-trial uncontrolled variation in the stimuli present characteristic of the lighted box. With more trials, more of these stimuli get connected to the crossing response. As this happens, the probability of crossing within the specified time (median-latency criterion) will rise monotonically, according to Postulate 3.

During the dark cycle many responses can occur, but there is no consistent "last" response. From trial to trial, stimuli characteristic of the dark box as well as those common to both the light and dark box can become conditioned to either a crossing response or some response incompatible with crossing. Consequently, the postulates predict no progressive or cumulative association of these stimuli and responses. However, the Postremity Principle does tell us that during the dark period, the stimuli characteristic of the lighted box must remain faithfully connected to the crossing response.

The consequences of this theoretical analysis are in agreement with the observation of a rising probability curve of learning for the L-to-D condition.

Dark-to-light condition.—Interchange the words "dark" and "light" in the above analysis and it applies verbatim to the L-to-D condition. The empirical consequences are deduced to be the same but are not observed to be so in Fig. 1. We con-

clude that a logical relation has been shown between theory and data, and that in part the data deny the theory. Now the possibility is to be considered that alternative explanations of these results can be made without altering the basic postulates.

One reason likely to be offered by contiguity theorists for the failure of rats to learn to run *into* a light assumes the presence of a "freezing" response to light or some other response incompatible with the to-be-learned crossing response. A "freezing" response, if it exists, would be immediately preceded by dark-box stimuli and become associated with them according to Postulate 1. The "freezing" response would follow the crossing response, hence the Postremity Principle predicts that it would be dominant. This is an adequate alternative explanation if a good case can be made for the "freezing" response. Such a response must be inferred. It is not easily observed to light, as it is to other strong stimuli such as shock or sound. Some indirect evidence for it consists of the fact that there is lowered activity level in the light (**4, 9**) and that there is some suggestion of a progressive *weakening* of the response of crossing into the light in the present experiment. More evidence than this will be required to make the inference secure, however. What is needed here is a program of research on the molecular analysis of a rat's unlearned responses to the onset of a light. The crucialness of the present experiment for theory depends on the outcome of such a program. We have some experiments on this problem underway.

A second reason for the failure of learning when the response turns on a light may be that the light merely adds stimuli and does not remove those of the dark box. There could be no last or "postreme" response, then, as far as the dark stimuli are concerned. Both crossing and not-crossing responses could occur in their presence at different times and become associated in a nonsystematic fashion. Consequently the postulate set would predict the absence of any cumulative learning effect.

This, too, is an adequate alternative explanation provided that we can accept a specification of stimuli that gives the dark box no unique stimuli in our set-up. The definition of stimulus accompanying the Voeks-Guthrie postulate set uses a criterion of afferent impulses in the CNS. If the rat eye is like the rabbit eye in displaying an "off-response" (**1**, p. 281ff.), there is some physiological basis for assuming stimulus elements present in the dark not present in the light. But whatever the physiological picture may turn out to be, there is other evidence of a behavioral nature that a stimulus specification of light as mere addition of stimuli leads to trouble when certain facts of discrimination are considered. We know that rats in a Skinner box can associate an R (acquisition of a lever press) with the cues provided by the dark box, and associate a not-R (extinction of lever press) with the cues of the same box lighted (**6**, p. 167ff.). This would not be possible if the dark cues were all present during the light. Postulate 2 would lead us to the contrary-to-fact conclusion that the dark cues, unprotected by the light, would become associated with the not-R every time an extinction trial was run. It is concluded that this is the weaker of the two alternative explanations.

These alternatives are not regarded as exhaustive of the reasonable possibilities. Others may undoubtedly be found which are testable and suggestive of further research. A highly probable alternative, or perhaps the

best contiguity postulate set, for these data has not yet been uncovered.

Summary

The fact that rats will learn a response that turns off a light but not one that turns on a light is offered as a test of two theories of the mechanism of learning. Hull's reinforcement theory adequately accounts for the data. Guthrie's contiguity theory as formalized by Voeks is not directly confirmed. The crucialness of the data for Guthrie's theory is seen to be dependent upon the outcome of further research on the nature of a rat's unlearned response to light.

(Received August 25, 1952)

References

1. Bartley, S. H. *Vision.* New York: D. Van Nostrand, 1941.
2. Guthrie, E. R. Conditioning: A theory of learning in terms of stimulus, response, and association. *Forty-first Yearbook of the National Society for the Study of Education.* Bloomington, Ill.: Public School Publishing Co., 1942.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
4. Hunt, J. McV., & Schlosberg, H. The influence of illumination upon general activity in normal, blinded, and castrated male white rats. *J. comp. Psychol.*, 1939, **28**, 285–298.
5. Keller, F. S., & Oberlin, K. W. A simple technique for measuring light-dark preference in the white rat. *J. genet. Psychol.*, 1942, **61**, 163–166.
6. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
7. Voeks, W. V. Stimulus stability and the acquisition of S–R connections. Unpublished doctor's thesis, Yale Univer., 1947.
8. Voeks, W. V. Formalization and clarification of a theory of learning. *J. Psychol.*, 1950, **30**, 341–362.
9. Zeaman, D., & House, B. J. Response latency at zero drive after varying numbers of reinforcements. *J. exp. Psychol.*, 1950, **40**, 570–583.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

RETENTION OF SERIAL NONSENSE SYLLABLES AS A FUNCTION OF REST-INTERVAL RESPONDING RATE AND MEANINGFULNESS[1]

E. JAMES ARCHER

Northwestern University[2]

In the area of verbal learning there has appeared somewhat capriciously a phenomenon known as reminiscence. Broadly defined, "*Reminiscence is an improvement in performance as shown by some measure of ability to recall at some time after the original practice, without (any) intervening practice*" (2, p. 314). Some studies (3, 5, 7, 8, 9, 10, 19, 20) have reported reminiscence while others (3, 4, 11, 14, 15, 16, 17) having very similar conditions have failed to observe it. As with reliable phenomena, it is expected that reminiscence either occurs or does not occur depending upon the existence or nonexistence of certain conditions. The search for these conditions leading to a body of predictive laws has not been very fruitful. There is in fact a remarkable lack of known variables which affect reminiscence in any systematic fashion. The few apparent dimensions which we know of have been examined to so limited an extent as to raise serious questions about their pertinence for reminiscence.

Without considering these dimensions in detail as was done by Buxton (2) in 1943 (little of theoretical consequence has been added since) we can describe the apparently necessary although not sufficient conditions for reminiscence. Reminiscence may be observed if 12 or 16 nonsense syllables are learned by the serial anticipation method either by spelling without correction or by pronouncing with correction to a moderate degree of mastery and if the rest interval is filled with color naming, color association, or light reading. When observed, reminiscence appears to increase to a maximum after about 2 min. and then decrease to zero between 5 and 10 min.

Several theories of reminiscence have been proposed, e.g., Hull *et al.* (12), but none can handle the many failures to find reminiscence. A study by Underwood (18) suggests a new explanation. The effects of two rest-interval activities were compared in a massed- versus distributed-practice study and it was observed that *S* made more errors per trial when color naming was used as the rest-interval activity than when symbol cancellation was used. Furthermore, following the introduction of the two activities on the practice day there was some evidence that the color-naming group learned faster than the symbol-cancellation group. Clearly the different rest-interval activities affected the rates of responding differentially. Thus it may be concluded that color naming either increased responsivity (tendency to respond) of *S* or symbol cancellation decreased responsivity.

The most obvious common dimension (from the point of view of *S*'s behavior) in these two rather dissimilar activities is the ratio of responses made to stimuli perceived (R/S). In color naming *S* is instructed to say

[1] This article is based in part on a dissertation submitted to the Graduate School, Northwestern University. The writer is indebted to Dr. Benton J. Underwood for his many helpful suggestions.

[2] Now at the University of Wisconsin.

aloud the name of *each* patch of color; the R/S ratio is 1/1. In the symbol-cancellation task employed by Underwood *S* was instructed to draw a short line through certain symbols, e.g., %, $, #, ", on a mimeographed sheet. The *S* cancelled (responded to) 9 of 30 symbols; therefore *S* had an R/S ratio of 9/30 or 3/10.

If this dimension is pertinent to reminiscence, it should be possible systematically to vary the R/S ratio, using the same material throughout, and to vary the responsivity of *S*. If, furthermore, on a recall trial following a short filled rest period there are a few subthreshold syllables, i.e., items which have been learned nearly to the point of correct anticipation, the number of such borderline items recalled should vary directly as the level of induced responsivity. It was further reasoned that the effectiveness of a rest-interval activity for inducing responsivity will depend upon the similarity of the learning task and the rest-interval activity. Accordingly a rest-interval activity was selected having the highest degree of similarity to the learning task possible while keeping retroactive inhibition at a minimum.

Another variable incorporated into the design was that of meaningfulness.[3] Several lines of evidence suggest that responsivity may be better induced in low-meaningful material and therefore reminiscence would be a function of meaningfulness. Furthermore if retention is intrinsically better for high-meaningful material, i.e., less dependent upon number of reinforcements, the time course of reminiscence may be altered.

To summarize, the purpose of the present experiment was to test the effect of rest-interval responding rate and meaningfulness on short-interval retention of serial nonsense syllables. If reminiscence varies as a function of rest-interval responding, some of the inconsistencies in the literature may be accounted for.

Procedure

Subjects.—A total of 72 Northwestern University undergraduates (36 men and 36 women) participated in this experiment for five consecutive days (one practice, four experimental). All *S*s learned the same list (medium association value) under the same conditions (all R/S ratios) on the first day (practice). On the basis of trials required to learn to a criterion, *S* was assigned to one of six groups having nearly equal means and variances.

Syllable lists.—On each of the four experimental days *S*s learned a different 13-item serial nonsense-syllable list. The first syllable served as a start cue. These lists were typed on white vellum tapes and presented on a modified Hull-type memory drum at a 2-sec. rate with 4-sec. rest between trials.

There were two sets of four lists. One set (high meaningful material) was made of syllables having 87%–100% association value (6). The other set (low-meaningful material) was made of syllables having 0%–13% association value. Syllables containing *Y* were never used. Intra-list similarity was minimal in all lists since a different consonant started each of 13 syllables in a list. The minimum number of six consonants both started and ended syllables; all vowels were used about equally often.

Three of the six groups ($N = 36$) learned high-meaningful lists and the other three learned low-meaningful material. Each day *S* learned a list until he was able to anticipate correctly 7 of 12 syllables. This criterion was followed with one of four rest intervals after which learning was resumed and continued until *S* correctly anticipated the entire list on a single trial. One rest interval was the usual 4-sec. period between trials and was therefore the control "no-rest" condition. The other intervals of 2, 5, and 10 min. were filled with a rest-interval activity.

Rest-interval activity.—A rest-interval activity having high similarity to the learning task while producing minimal retroactive inhibition was selected to induce most effectively controlled responsivity. Three-digit numbers were typed on tapes one-half as wide and four times as long

[3] After this study was completed, it was learned that Melton (13) and Noble (16) had also measured rate of learning as related to meaningfulness. Melton's paper has not been read by the present author, but it is referred to by Noble.

as the syllable tapes. These numbers were spaced in serial order of 13 to a unit followed by two blank spaces so that number tapes had the same physical characteristics as syllable tapes except that numbers replaced letters. Four units of 13 three-digit numbers were used to prevent any learning of the rest-interval activity; *S* was also specifically instructed *not* to try learning the numbers. Furthermore, to prevent learning, a different number tape was used each day.

Each of the three groups of *S*s within a level of meaningfulness had a different responses-made-to-stimuli-perceived ratio (R/S) during the rest intervals. This ratio was controlled by instructing *S* to read aloud those numbers which were marked with an asterisk to the left of a certain number of numbers per unit. Consequently there were three sets (corresponding to three R/S ratios) of four tapes (one for each day), each tape having four units of 13 three-digit numbers.

One set of four tapes had an asterisk to the left of three numbers per unit (an R/S ratio of 3/13). A second set had an additional four numbers so marked (an R/S ratio of 7/13). A third set had all but one number per unit so marked (an R/S ratio of 12/13). On the practice day a tape of three-unit length (one unit for each R/S ratio) was used.

The *S* was told his reaction time to the numbers was being recorded. This was done to maintain his interest in the rest-interval activity. Most *S*s thought their reaction times were being recorded and therefore, as instructed, avoided rehearsal of the learning tasks "so as to not slow down their reaction times."

Experimental design.—A modified greco-latin analysis of variance design (1) was employed to obtain maximum information from the data. The modification was the addition of another counterbalanced internal variable yielding what might be described as a seven-classification arabic-greco-latin repeated measures analysis of variance. Figure 1 is a diagram of this design, which essentially shows 72 *S*s divided on the basis of *meaningfulness* into two large groups. These are further divided into three levels of *R/S ratios*. The final horizontal classification is based on the 12 *S*s in each R/S ratio. The single vertical classification is by *experimental days* of which there are four. Three internal variables have been counterbalanced in 18 4 × 4 arabic-greco-latin squares.

Each cell in these squares has three designators—one each for the *number list* (Arabic numeral), *syllable list* (Greek letter), and *length of rest interval* (Latin letter)—for the combination used on a particular day for a particular *S*. The chief advantage of this design is that total variance may be analyzed to test the influence of all main effects and several important interactions.

Meaning-fulness	R/S	Ss	Experimental Day I	II	III	IV
High	12/13	1	2 α B	1 δ C	3 β D	4 γ A
		2	3 δ A	4 α D	2 γ C	1 β B
		3	1 γ D	2 β A	4 δ B	3 α C
		4	4 β C	3 γ B	1 α A	2 δ D
		⋮ 12				
	7/13	13 ⋮ 24				
	3/13	25 ⋮ 36				
Low	12/13	37 ⋮ 48				
	7/13	49 ⋮ 60				
	3/13	61 ⋮ 72				

Fig. 1. A schematic representation of the experimental design. The arabic-greco-latin square for the first four *S*s is shown in detail.

Results

Comparability of groups.—Evidence for initial equality of groups is provided by an analysis of variance of total trials to learn on practice day. The analysis shows a significant restriction in the variance of the six groups ($1/F = 41.14$; with 66 and 5 *df*, a $1/F$ of 9.24 is needed for $p = .02$) and indicates that the groups were reasonably equal in ability initially.

Evidence for the efficacy of matching groups is provided by the correlation between the matching variable (trials to learn on the practice day) and total trials to learn on all experi-

TABLE 1

ANALYSIS OF VARIANCE OF CODED CORRECT DIFFERENCE SCORES

Source of Variation	*df*	Mean Square	*F*
Meaningfulness	1	0.22	
Responding rate	2	10.53	3.35*
Meaning. × Resp.	2	0.15	
Ss/Responding rate	66	3.14	1.08
Rest interval	3	50.94	17.57***
Rest × Meaning.	3	8.23	2.84
Rest × Resp.	6	2.21	
Day	3	6.21	2.14
Day × Meaning.	3	8.79	3.03
Day × Resp.	6	4.27	1.47
Syllable lists (Pooled)	6		
Syll. × Resp. (Pooled)	12		
Number lists (Pooled)	9		
Number × Meaning. (Pooled)	9		
Rest × Meaning. × Resp.	12		
Day × Meaning. × Resp.			
Ss × Day/Resp. (Pooled)	144	2.90	
Total	287		

* Significant at .05 level.
*** Significant at .001 level.

mental days combined. This correlation was .57 ($\sigma_r = .12$).

Reminiscence.—In this study a performance measure on the recall trial was subtracted from the comparable measure on the criterion trial and ten was added to this difference to eliminate negative values. These values, coded difference scores, measure recall performance as a function of actual performance at the criterion. Table 1 presents an analysis of coded correct difference scores. The error term for testing Meaningfulness, Responding rate, and Meaning × Responding rate is *Ss*/Responding rate. The effect of Meaningfulness was not significant, i.e., high-meaningful nonsense syllables were not retained any better than the low-meaningful nonsense syllables. The effect of Responding rate was just significant ($.05 > p > .04$) but in an unexpected manner. Fewer correct responses were given after the 7/13 number-naming activity than after 3/13 or 12/13, which were about equally effective.

For the remaining variance components the *Ss* × Day/Responding rate is the proper error term. The pooled error term, however, had significant heterogeneity (chi square = 13.58; $.02 > p > .01$) which could not be eliminated with a square root or log transformation.

Since heterogeneity of variance was present in this second error term, the two interactions (Rest × Meaning, and Day × Meaning) were not considered significant as the corresponding *F* ratios just exceeded the .05-level requirement. Consistent with this rejection of significance was the absence of a systematic change in the means of the corresponding classification subgroups.

In spite of heterogeneity of variance, length of rest interval was considered a significant source of variance ($F = 17.56$; $p < .001$). The effect of this variable, as may be seen in Fig. 2, is that recall decreases as rest interval increases. No initial increase was

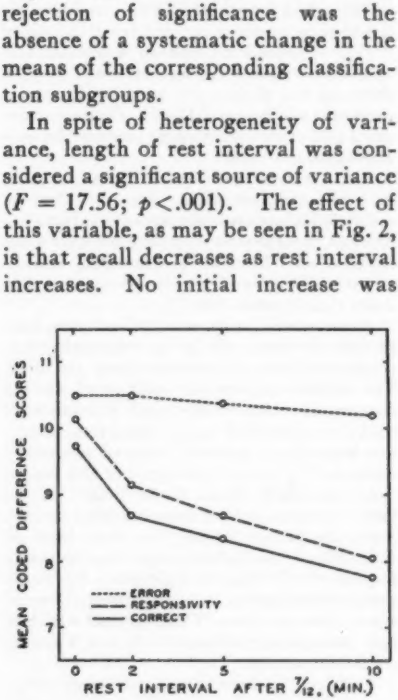


FIG. 2. Mean coded difference scores as a function of length of rest interval. Only the slopes and not the ordinates of curves are comparable.

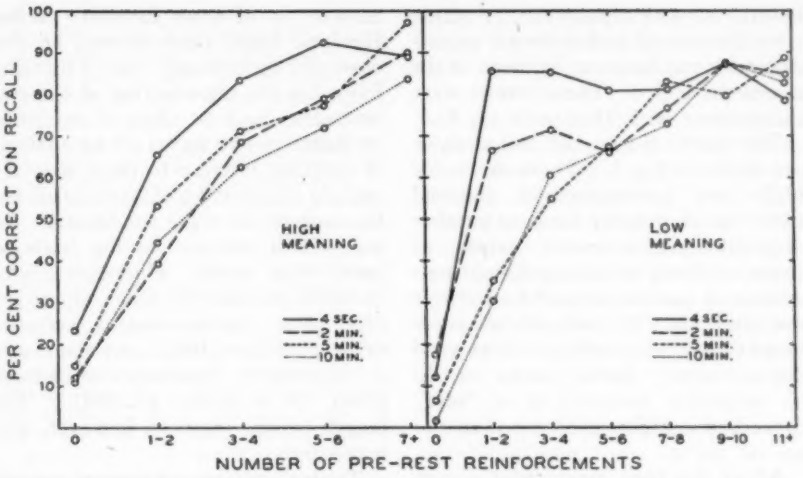


FIG. 3. Percentage correct on recall as a function of length of rest interval and number of times item was correctly anticipated before rest interval for high- and low-meaningful material

observed even when the curves were plotted for each of the six responding-rate groups. Clearly there is no evidence for reminiscence. In fact after 2 min., the point at which others have found reminiscence at a maximum, a significant degree of forgetting was obtained ($t = 3.90; p < .001$). It should be noted that the values plotted in Fig. 2 indicate the change in responses and not the absolute values. That is, although the error curve is above that for coded correct difference scores, more correct responses were made than errors.

Responsivity.—It may be argued that the preceding analysis was not a fair test of the effectiveness of the rest-interval activity for inducing a controlled set to respond, i.e., for changing the responsiveness of *S*. A better test might be to compare the change in total responses, both correct and incorrect, between the corresponding criterion and recall trials. Coded difference scores were computed and analyzed as before. None of the variables, except length of rest interval, was significant. Since the "error" curve is very nearly a horizontal straight line, the form of the "responsivity" curve as shown in Fig. 2 is almost identical with the "correct" curve except for a vertical displacement.

Errors.—The analysis of the coded overt-error difference scores which are analogous to the preceding scores but based upon incorrect responses, does not have any direct relevance to the problem under consideration. However, it does have some theoretical interest. None of the variables was significant. Even the length of the rest interval was without effect, that is, there was neither an increase nor a decrease in errors as a function of the length of the rest interval.

An analysis of pre- and postrest error ratios (errors per trial) indicated that meaningfulness, responding rate, and length of rest interval were ineffective.

Retention of individual items.—Mean

failures at recall at each serial position were determined and there was no evidence for reminiscence of items in the middle of a list. These results were inconsistent with Hovland's (7, 8).

The results of a second item analysis are shown in Fig. 3. These curves for high- and low-meaningful material show the relationship between number of prerest reinforcements (number of times correctly anticipated) and percentage of success on recall for the four rest intervals. In these sets of curves forgetting is indicated by a downward displacement. Reminiscence would be indirectly indicated if a "rest" curve had a higher ordinate than the no-rest curve.

All of the high meaningful curves appear to have essentially the same negatively accelerated form. The only differences appear to be in vertical displacement; except for the 2-min. curve this change in retention is orderly.

On the other hand both form and displacement are different in the low-meaningful curves. Having one or more reinforcements makes it very probable (85%) that the item will be given correctly on the very next trial if there is no rest. It is considerably less probable (60%) that an item will be recalled after 2 min. if it has but one or two reinforcements. There appears to be a complex interaction between degree of learning, retention interval, and meaningfulness. For the high-meaningful material the rate of forgetting appears to be about constant for all degrees of learning. This is not true of low-meaningful material for which the rate of forgetting is much greater for a low degree of learning than for a high degree.

Trials to learn.—Although not a proper test of reminiscence as previously defined, the effect of length of rest interval on total trials to learn to mastery is of some interest. When Hovland found reminiscence, he also observed an increased rate of learning following the introduction of the rest interval. Such an effect of rest could be shown in two ways: (*a*) an analysis of variance of trials to reach mastery and (*b*) the plotting of learning curves for each of the eight conditions. An analysis of variance of log trials to learn was made. High-meaningful material was learned much more rapidly than low-meaningful material ($F = 44.69$; $p < .001$); and there was a significant learning-how-to-learn effect ($F = 33.56$, $p < .001$). The length of rest interval, however, was ineffective.

Trials to criteria curves were plotted for the two levels of meaningfulness. As would be expected from the results of the analysis of variance, the two sets of four curves had distinctly different slopes. More important than that, however, was the negative finding of no change in slope as a result of introduction of the rest interval. All four curves at one level of meaningfulness had essentially the same shape.

Discussion

The number of responses made compared to stimuli perceived (R/S ratio) was expected to account for the "now-you-see-it-now-you-don't" (2, p. 337) characteristic of reminiscence. Instead this variable proved to be ineffective.

There are at least two possible explanations as to why the different rest-interval activities were ineffective and reminiscence was not observed. (*a*) Although *S* was instructed to read aloud only those numbers preceded by an asterisk, it would not have been difficult for him to read *covertly* those numbers not so designated. If *S* habitually made many covert responses, there would be practically no difference between the responding rates. (*b*) The rest-interval activity used in this study was antagonistic to the learning material in that during learning *S* was to anticipate the next syllable and say it aloud before it appeared, whereas during the number-naming activity *S* was to read aloud only those numbers appearing in the aperture

and to avoid trying to guess what the next numbers would be. In the former case *S* must learn to anticipate and in the latter case *S* must learn not to anticipate. On the recall trial, during which anticipation is essential to success, the number-naming activity may have produced sufficient negative transfer to produce forgetting. It is of significance that unlike Underwood (18) we failed to change the postrest responsivity of *S* and therefore the original hypothesis that high and low responsivity could account for reminiscence and forgetting respectively has not been tested. Perhaps the common dimension of responding rate was not the pertinent one for altering responsivity.

Since there was a very wide difference in rate of learning, it is interesting that recall after fast learning (high meaningful) was not better than after slow (low meaningful). If a relationship obtains between rate of learning and recall, we would expect a negative correlation between trials to criterion and coded correct difference scores. For faster learning (few trials) the projected slope of performance should be higher (higher coded correct difference scores). The reverse should hold for slower learning. Of eight correlations (one for each rest interval at each level of meaningfulness) between trials to criterion and coded correct difference score, only one was significantly greater than zero. Rate of learning and recall, over the intervals tested, do not appear to be correlated.

Another test of this lack of correlation was made by comparing the coded correct difference scores of the 18 *fastest Ss* who had learned high-meaningful material with the scores of the 18 *slowest Ss* who had learned low-meaningful material. The retention interval was 4 sec. (a trial-to-trial test). There was no significant difference ($F = 1.09$) between recall measures of these two groups although there was a very significant difference in rate of learning ($F = 89.20$, $p < .001$). This lack of correlation argues against the likelihood of a significant meaning × reminiscence interaction.

Summary

An experiment was performed in which retention of serial nonsense syllables was measured as a function of length of rest interval, meaningfulness of syllables, and rest-interval responding rate.

Six groups of 12 *Ss*, each nearly equal in learning ability, served in this experiment. The groups corresponded to the six possible combinations of two levels of meaningfulness and three levels of rest-interval activity. On each of four consecutive days *S* learned a different 12-syllable list until he correctly anticipated seven on a single trial. He then had one of four rest intervals, 4 sec., 2 min., 5 min., or 10 min., the last three being filled with a rest-interval activity. The lengths of rest interval, syllable lists, and number lists, were counterbalanced for all four days in an arabic-greco-latin design. Following the rest interval *S* resumed learning and continued until he correctly anticipated the entire list.

The basic findings may be summarized as follows:

1. Recall decreased as the rest interval increased and was unrelated to meaningfulness or rest-interval activity.
2. Responsivity also decreased as the interval increased and was not affected by meaningfulness or rest-interval activity.
3. The number of errors per trial before and after the rest interval was unrelated to meaningfulness, rest-interval activity, or length of rest interval. As *S* became more practiced, he made fewer postrest errors.
4. There appeared to be a complex interaction between meaningfulness, prerest reinforcements, and length of rest interval as measured by recall.
5. Meaningfulness and stage of practice were the only variables which affected the total number of trials to attain mastery.

The results obtained do not support the hypothesis that reminiscence would vary as a function of rest-interval responding rate or of meaningfulness. Possible explanations of these results were offered.

(Received September 2, 1952)

References

1. Archer, E. J. Some greco-latin analysis of variance designs for learning studies. *Psychol. Bull.*, 1952, **49**, 521–537.
2. Buxton, C. E. The status of research in reminiscence. *Psychol. Bull.*, 1943, **40**, 313–340.
3. Buxton, C. E. Repetition of two basic experiments on reminiscence in serial verbal learning. *J. exp. Psychol.*, 1949, **39**, 676–682.
4. Buxton, C. E., & Bakan, M. B. Correction vs. non-correction learning techniques as related to reminiscence in serial anticipation learning. *J. exp. Psychol.*, 1949, **39**, 338–341.
5. Buxton, C. E., & Ross, H. V. Relationship between reminiscence and type of learning technique in serial anticipation learning. *J. exp. Psychol.*, 1949, **39**, 41–46.
6. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
7. Hovland, C. I. Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and by distributed practice. *J. exp. Psychol.*, 1938, **22**, 201–224.
8. Hovland, C. I. Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **22**, 338–353.
9. Hovland, C. I. Experimental studies in rote-learning theory. IV. Comparison of reminiscence in serial and paired-associate learning. *J. exp. Psychol.*, 1939, **24**, 466–484.
10. Hovland, C. I. Experimental studies in rote-learning theory. VI. Comparison of retention following learning to same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, **26**, 568–587.
11. Hovland, C. I., & Kurtz, K. H. Experimental studies in rote-learning theory. IX. Influence of work-decrement factors on verbal learning. *J. exp. Psychol.*, 1951, **42**, 265–272.
12. Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., & Fitch, F. B. *Mathematico-deductive theory of rote learning.* New Haven: Yale Univer. Press, 1940.
13. Melton, A. W. A comparative study of the materials employed in experimental investigations of memory. Unpublished M.A. thesis, Yale Univer., 1929.
14. Melton, A. W., & Stone, G. R. The retention of serial lists of adjectives over short time-intervals with varying rates of presentation. *J. exp. Psychol.*, 1942, **30**, 295–310.
15. Noble, C. E. Absence of reminiscence in the serial rote learning of adjectives. *J. exp. Psychol.*, 1950, **40**, 622–631.
16. Noble, C. E. The role of stimulus meaning (*m*) in serial verbal learning. *J. exp. Psychol.*, 1952, **43**, 437–446.
17. Shipley, W. C. The effect of a short rest pause on retention in rote series of different lengths. *J. gen. Psychol.*, 1939, **21**, 99–117.
18. Underwood, B. J. Studies of distributed practice: VI. The influence of rest-interval activity in serial learning. *J. exp. Psychol.*, 1952, **43**, 329–340.
19. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 4 (Whole No. 220).
20. Withey, S., Buxton, C. E., & Elkin, A. Control of rest interval activities in experiments on reminiscence in serial verbal learning. *J. exp. Psychol.*, 1949, **39**, 173–176.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

STUDIES OF DISTRIBUTED PRACTICE: X. THE INFLUENCE OF INTRALIST SIMILARITY ON LEARNING AND RETENTION OF SERIAL ADJECTIVE LISTS

BENTON J. UNDERWOOD[1]

Northwestern University[2]

The major purpose of this study was to discover the influence of intralist similarity and intertrial interval on retention of serial adjectives. With regard to retention as a function of intralist similarity, several studies (**4, 6, 7, 8**) indicate that while intralist similarity may influence rate of learning, it has no influence on recall taken 24 hr. following learning. As yet, however, no tests of retention have been made for serial adjectives of varying levels of intralist similarity.

The data are conflicting concerning retention of rote-learned materials following acquisition with varying intertrial intervals. With serially learned nonsense syllables, Hovland (**3**) found better retention after spaced practice than after massed practice. We have found quite the opposite with comparable materials (**6**). Retention of paired adjectives may (**4**) or may not (**8**) be enhanced by distributed practice depending, we believe, upon degree of learning before the retention interval. Although degree of learning has not clearly been shown to be the critical variable, some evidence (summarized in **8**) suggests that with high degrees of learning massed practice gives better recall, and with low degrees of learning spaced practice gives better recall. In any event, data available at the present time do not justify generalizations concerning retention following learning by massed and distributed practice. Since there are no results on this variable for serial adjectives, the present experiments will fill this gap.

Previous experiments (**5, 9**) on learning of serial adjectives as a function of massed vs. distributed practice have shown small but consistent differences in favor of faster learning with distribution. The present experiments will give additional data on this relationship, for obviously, in order to measure retention following different intertrial intervals, learning rates under these intervals will also be available.

Procedure: Exp. I, II, and III

General.—The first data to be reported are based on three experiments. Two additional experiments make up part of the report but will not be described until after the results of the first three studies are given. The first three experiments are differentiated only by degree of intralist similarity of the adjective lists. In Exp. I, intralist similarity is low; in Exp. II, medium; and in Exp. III, intralist similarity is high. Each experiment had three different conditions, these conditions varying in length of intertrial rest interval. These intervals were 2 sec. (massed), 30 sec., and 60 sec.

Lists.—Since there were three conditions in each experiment, and since a given *S* served in all three conditions, three different experimental lists were required for each experiment. Each serial list consisted of 14 two-syllable adjectives, but since the first word was used only as an anticipatory cue, *S* learned only 13 items. All adjectives were taken from Haagen (**2**). In constructing the lists, the 14 items to be used were formed into five sets in which four sets consisted of three words and one set of two words. In varying similarity, synonymity of items within

[1] R. L. Morgan and E. J. Archer supervised the gathering of the data; Mr. Archer and Jack Richardson are largely responsible for the statistical analysis.

[2] This work was done under Contract N7onr-45008, Project NR 154-057, between Northwestern University and the Office of Naval Research.

sets was varied. That is, for low similarity the items within sets making up the lists would have low synonymity while for high similarity the items within sets had a high-synonymity rating. The lists, and sets within lists, are illustrated below for each level of similarity. Sets are enclosed in parentheses.

Exp. I: Low Similarity: (fickle, heedless, fitful) (urgent, crying, required) (sickly, bedfast, feeble) (complete, perfect, utter) (empty, yawning)

Exp. II: Medium Similarity: (anxious, distressed, troubled) (cautious, discreet, guarded) (swollen, bloated, puffy) (liquid, flowing, solvent) (spoken, talking)

Exp. III: High Similarity: (angry, enraged, wrathful) (complete, entire, total) (double, dual, twofold) (royal, regal, kingly) (liquid, fluid)

If one studies the above illustrations, it will be recognized that degree of similarity within sets increases from Exp. I through Exp. II. The method of increasing similarity, it will be realized, does not increase the number of potential interfering tendencies but increases the strength of those tendencies. In Haagen's scaling, .5 indicates highest similarity and 6.5 lowest. Averaging the synonymity ratings for the sets of each list for Exp. I gives means of 4.04, 3.80, and 4.01; for Exp. II the means are 2.40, 2.42, and 2.40; for Exp. III the means are 1.22, 1.31, and 1.22.

In ordering items for a list, of course, synonyms were never placed together but were scattered throughout the list. The individual familiarity ratings of all adjectives in a list averaged about the same for all lists for all three levels of similarity. Obvious associative cues (e.g., same prefix on two successive words) were kept at a minimum.

A single practice list, learned by all *S*s, was of medium similarity. All lists were presented on Hull-type drums at a 2-sec. rate, with the anticipation method of learning used throughout.

Specific conditions.—On the practice day *S* learned the practice list to seven correct responses on a single trial by massed practice (2 sec. between trials). Instructions for symbol cancellation, used to fill rest intervals between trials, were followed by continued learning of the list to one perfect trial with 30-sec. rest between each trial. After 5-min. rest the practice list was recalled and relearned.

On the experimental days *S* learned a list to one perfect recitation under the intertrial interval appropriate for the day. After 24 hr. the list was recalled and relearned by massed practice. On the first three experimental days *S* learned a new list each day which was in turn recalled the following day. On the last of the four experimental days only recall and relearning of the list learned the previous day were required. Thus, *S* served in five sessions, the first being a practice session. The four experimental sessions came on four successive days in order to meet the 24-hr. retention interval for each list.

A total of 36 undergraduate students was used in each experiment. Conditions of intertrial rest and lists were completely counterbalanced against practice effects. The method for statistical analysis of such designs has been reported elsewhere by Archer (1).

Results: Exp. I, II, and III

Practice list.—The mean number of trials to learn the practice list to one perfect recitation was 27.36, 29.72, and 26.25 for Exp. I, II, and III, respectively. F is less than one. The product-moment correlation between trials to learn practice list and trials to learn all three experimental lists for all three experiments combined is $.60 \pm .10$. The mean number of errors per trial in learning the practice list was 1.83, 1.54, and 1.74. F is 1.32, with an F of 3.09 needed for significance at the .05 level of confidence. The correlation between errors

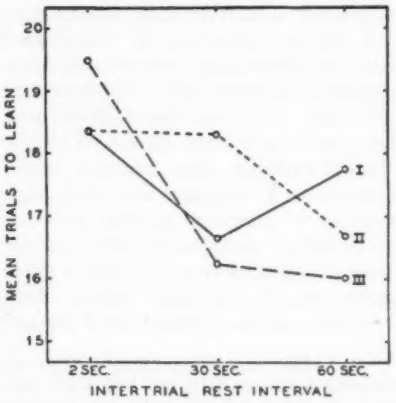


Fig. 1. Learning as a function of intertrial interval and intralist similarity. Exp. I had lists of low similarity; Exp. II medium; and Exp. III, high intralist similarity.

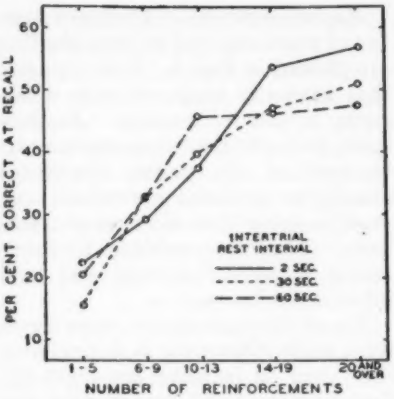


FIG. 2. Retention as a function of intertrial rest with number of reinforcements (correct responses) during learning held constant

on the practice day and errors on experimental days was .64 ± .10. The three groups of *Ss* may be considered comparable.

Experimental lists.—The mean number of trials to learn the lists of varying similarity under three intertrial rest conditions is shown in Fig. 1. The statistical analysis of the distributions on which these means are based has shown: (*a*) Similarity as varied here is not an effective variable, *F* being less than one. (*b*) Intertrial rest is a significant source of variance. The *F* for intertrial rest is 4.48. With 2 and 200 *df*, *F* is 3.04 at .05 level of confidence, and 4.71 at the .01 level. (*c*) Interaction between intertrial rest and intralist similarity is far from significant (*F* is 1.33).

Mean errors per trial did not vary as a function of intertrial interval. Combining all three intertrial rest conditions for each experiment, the mean number of errors per trial was 1.22, 1.21, and 1.45, for Exp. I, II, and III, respectively. The distributions on which these means are based do not differ significantly.

Recall and relearning.—Analysis of the 24-hr. recall shows no significant source of variance. The trend for the raw recall scores was for better recall for high-similarity lists than for low-similarity lists. In order to evaluate the influence of intertrial rest on retention we have made an item analysis of original learning by grouping items having comparable number of reinforcements (correct anticipations) and then determining percentage correct at recall for each grouping. The result of this analysis is shown in Fig. 2. While the differences are small, it may be noted that recall is better following massed practice than following spaced practice for items receiving a large number of reinforcements. The situation is reversed for items with a moderate number of reinforcements. As summarized elsewhere (**8**), comparable trends have been noted in several other studies.

Neither interlist similarity nor intertrial interval was effective during relearning.

All findings for learning, recall, and relearning indicate that similarity as manipulated here was an ineffective variable. The only clear positive finding of the three experiments was that distributed practice produced faster learning than massed practice. Even here the level of significance was not high and confirms previous findings (**5, 9**) that with serial adjectives distributed practice produces small but consistent amounts of facilitation.

In view of the fact that similarity did not influence rate of learning, we have no adequate test of retention as a function of intralist similarity. Therefore, we have run two additional experiments to obtain more definitive data on this matter. The purpose was to obtain lists varying in similarity which would produce differences in rate of learning.

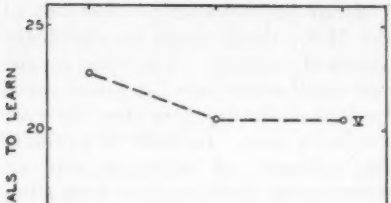


FIG. 3. Learning as a function of intertrial interval and interlist similarity. Experiment IV had very low intralist similarity; Exp. V had very high intralist similarity.

PROCEDURE: EXP. IV AND V

The only difference between these two experiments and those reported above was in lists. In Exp. IV the lists had very low similarity. No known meaningful similarity existed between items within a list and all apparent instances of formal similarity were eliminated. In Exp. V, on the other hand, lists were constructed to have high similarity by increasing the number of synonymous words within a list. The 14-item lists were made with three sets of highly synonymous adjectives, one set containing four synonyms and two sets containing five. This is to be contrasted with Exp. III where we used four sets of three synonyms and one set of two. In short, all details of the experiments were the same as those reported above except that Exp. IV had lists constructed to have lower intralist similarity than Exp. I, and Exp. V had lists constructed to have higher intralist similarity than Exp. III.

RESULTS: EXP. IV AND V

Practice list.—The mean number of trials required to learn the practice list was 28.61 for Exp. IV, and 27.31 for Exp. V. The corresponding mean number of errors per trial was 1.36 and 1.40. Since these means for either measure do not differ significantly, we may consider the two groups of *Ss* equivalent.

Experimental lists.—The mean number of trials required to learn the lists are plotted in Fig. 3. It is apparent that similarity produced wide differences in rate of learning. Furthermore, as in the first three experiments reported in this paper, distributed conditions produced somewhat more rapid learning than did massed conditions. There is no evidence for interaction between intertrial rest and intralist similarity.

For all three conditions, more errors were made in learning high-similarity lists than in learning low-similarity lists but these differences were not great enough to be reliable statistically.

Retention.—Intertrial interval during learning produced no significant differences in recall. However, differences in recall attributable to similarity were very large, with the better retention occurring for the high-similarity lists. For all three conditions combined, the mean number of items recalled from low-similarity lists was 3.30; for the high-similarity items the corresponding value was 5.00. The *F*

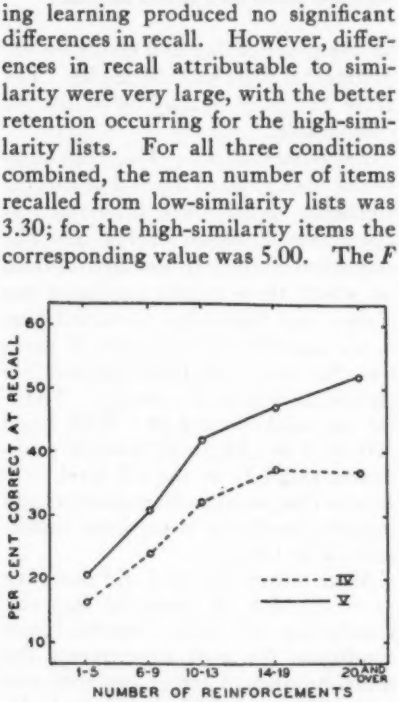


FIG. 4. Retention as a function of intralist similarity with number of reinforcements during learning held constant. Experiment IV had low similarity; Exp. V, high.

is 11.35, with an F of 7.01 needed for the .01 level with 1 and 70 *df*.

These retention results as a function of similarity must be evaluated cautiously. Since the number of trials required to learn the two sets of lists was widely different, it is apparent that mean frequency of reinforcements for items in the two lists was different. Therefore, differences at recall may reflect this variable and not similarity. Accordingly, an item analysis has been made in which number of reinforcements during learning is held constant for the two lists and percentage correct at recall determined. The results of this analysis are shown in Fig. 4. It can be seen that even with number of reinforcements held constant, recall of high-similarity items is superior to recall of low-similarity items. To make sure that Fig. 4 is not a result of bias in grouping reinforcements for learning, we have also analyzed recall for number of reinforcements from 1 through 20 separately. At nearly all points recall was better for high-similarity lists than for low-similarity lists. Clearly, the conclusion is that retention of serial adjectives with high intralist similarity is better than retention of serial adjectives of low intralist similarity.

Relearning required longer for high-similarity lists than for low, and more errors per trial were made in relearning the high-similarity lists than the low, but neither trend had statistical reliability.

Discussion

As in previous studies (**5, 9**) we have found that learning of serial adjectives is facilitated by distributed practice but this facilitation is not great. In all studies the level of statistical significance has not been high but always learning is faster with distribution than with massing. Furthermore, we have no clear evidence in any of these studies that there is interaction between intertrial interval and interlist similarity. As in the case of nonsense syllables (**6**) there is no basis for believing that difficulty as manipulated by intralist similarity is an important variable in the study of massed vs. distributed practice.

Retention as measured by recall scores has not varied as a function of intertrial interval in the five experiments reported here. Consistent with previous findings (**8**), however, is the trend noted from the item analyses of recall that heavily reinforced items are better recalled following massed practice than following distributed practice. Conversely, if the item is reinforced only a few times during learning, it will be better retained if learned under distributed practice than if learned under massed practice. It should be emphasized that these are only *trends* noted in the data reported here and in several previous experiments. To verify clearly these suggested principles, independent experiments will be necessary in which degree of learning is manipulated.

Finally, it should be pointed out that in all of our experiments on massing vs. distribution, whether using serial learning or paired-associate learning, or whether using meaningful material or nonsense syllables, no evidence has been found that intralist similarity is an important variable influencing recall. In the present study (Exp. IV and V) with degree of learning held constant, recall was consistently better for lists of high similarity than for those of low similarity. We do not believe that this finding is a contradiction to the above general statement. Rather, we believe that it is a reflection of the particular technique used to vary

similarity and as such represents an exception to the more general findings of previous experiments. It is, of course, important to know that there are ways in which similarity can be manipulated so that recall will be directly correlated with similarity, but the bulk of the evidence, in which similarity has been varied in many different ways, indicates the present finding to be an exception. In our high-similarity lists in the present experiments there were three clusters of synonyms. In constructing the lists two words from a given cluster were never allowed to be serially contiguous. With such a system it is possible to see how recall could be facilitated. Thus *S*, having just seen one word of a given cluster, may learn that the next word cannot be from the same cluster. Or, having seen two or three words from a given cluster during the first part of a trial, *S* might know with some certainty that among the last few words there would be one or two more words from that cluster.

In short, there are ways by which *S* could increase his recall for high-similarity lists but this increment could well be relevant only to the particular lists used here because of the method used in constructing them. Therefore, we do not feel that generalizations should be made from the present finding that high-similarity lists are recalled better than low-similarity lists.

Summary

Three experiments were performed to determine the influence of intralist similarity and intertrial interval on learning and retention of serial adjective lists. Similarity was manipulated by varying degree of synonymity of clusters within lists. Three intertrial rests of 2, 30, and 60 sec. were used for each experiment. The 36 *S*s in each experiment learned each list to one perfect trial following which retention was measured after 24 hr.

The results showed that the method of manipulating similarity did not produce differences in rate of learning. Learning was more rapid for the 30- and 60-sec. intertrial rest conditions than for the 2-sec. condition, thus confirming previous findings. No differences in retention were evident for either variable.

Two additional experiments were performed in an attempt to produce differences in rate of learning as a function of similarity. By manipulating number of items with a cluster of synonyms, wide differences in rate of learning were achieved. Again, distributed practice facilitated learning. The high-similarity lists were better recalled than low-similarity lists. Since several previous studies, using different materials and learning methods, and different techniques of manipulating similarity have shown little difference in recall as a function of interlist similarity, it was concluded that the better recall of the high-similarity lists in the present experiments was a function of the particular method of producing high similarity.

(Received September 2, 1952)

References

1. Archer, E. J. Some greco-latin analysis of variance designs for learning studies. *Psychol. Bull.*, 1952, **49**, 521–537.
2. Haagen, C. H. Synonymity, vividness, familiarity, and association-value ratings of 400 pairs of common adjectives. *J. Psychol.*, 1949, **30**, 185–200.
3. Hovland, C. I. Experimental studies in rote-learning theory: VI. Comparison of retention following learning to same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, **26**, 568–587.

4. Underwood, B. J. Studies of distributed practice: II. Learning and retention of paired-adjective lists with two levels of intra-list similarity. *J. exp. Psychol.*, 1951, **42**, 153–161.

5. Underwood, B. J. Studies of distributed practice: III. The influence of stage of practice in serial learning. *J. exp. Psychol.*, 1951, **42**, 291–295.

6. Underwood, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intra-list similarity. *J. exp. Psychol.*, 1952, **44**, 80–87.

7. Underwood, B. J. Studies of distributed practice: VIII. Learning and retention of paired nonsense syllables as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 133–142.

8. Underwood, B. J. Studies of distributed practice: IX. Learning and retention of paired adjective lists as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 143–149.

9. Underwood, B. J., & Goad, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–134.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

PERFORMANCE OF A MOTOR TASK AS A FUNCTION OF INTERPOLATION OF VARYING LENGTHS OF REST AT DIFFERENT POINTS IN ACQUISITION[1]

EUGENIA B. NORRIS

Perceptual and Motor Skills Research Laboratory, Human Resources Research Center[2]

In investigations of motor learning one of the factors that has been studied has been the effect of interpolation of rest periods during the acquisition of some motor skill. In general it has been found that when prerest practice has been massed to some degree, the interpolation of a rest leads to an increase in proficiency after the rest. This phenomenon has been called reminiscence. Another factor that has been studied in relation to reminiscence is the effect of length of rest. Bell (2) investigated the effects on amount of reminiscence in a pursuit rotor task of five lengths of rest varying between 10 min. and 30 hr. These rests were introduced after either 5 or 15 1-min. trials. He found that after five trials of prerest practice all lengths of rest led to approximately equal increases in score but that the effects of the longer rests were more lasting. After 15 min. of prerest practice the shorter rests led to an increase in score while the longer ones led to an initial postrest decrement followed by a rapid recovery.

Ammons (1), again using a pursuit rotor task, found that reminiscence as a function of amount of prerest practice increased to 8 min., then decreased to 17 min., reaching a level below that found after 3-min. prerest practice. In terms of the length of the interpolated rest, reminiscence rose quickly to a maximum after 5-min. rest and remained practically at this level through 6-hr. rest.

Kimble and Horenstein (4), investigating reminiscence with a pursuit rotor task as a function of length of interpolated rest, obtained results that indicated that reminiscence increased as a negatively accelerated function of length of rest and reached an asymptote around 600 sec.

In an experiment investigating reminiscence in pursuit rotor learning as a function of length of rest and amount of prerest practice, Irion (3) found that the amount of reminiscence at first increased and then decreased as amount of prerest practice increased from 2 to 40 trials. It was also shown that reminiscence was an increasing function of the length of rest through 5 min.

Several interesting facts should be noted about the above studies. First, only one task, the pursuit rotor, has been used. Second, none of the investigations ascertained the possible interaction of length of rest and amount of prerest practice. Third, with the exception of Ammons, the investigators introduced a series of rests, since the basic work schedule used was a trial of given length with a short rest between trials.

The present experiment was designed to find the effects on performance of a *continuous* two-hand motor task of (*a*) the interpolations of rest periods at different points in acquisition, (*b*) the length of a rest period, and (*c*) the possible interaction of these two variables.

PROCEDURE

Subjects.—The *S*s were basic trainee airmen at Lackland Air Force Base. They were assigned at random in units of four to one of seven groups until there were 100 in each of the groups.

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions and conclusions in this report are those of the author. They are not to be construed as necessarily reflecting the views or endorsement of the Department of the Air Force.

[2] Now at 3302nd Research and Development Group, Mather Air Force Base.

Apparatus.—The apparatus was an adaptation of the SAM Two-Hand Coordination Test (CM101B) which has been described elsewhere (5). Four such units were used and were run simultaneously. The adaptation consisted of modifying the test so that it would run continuously rather than with the standard 1-min. test and 15-sec. rest period. Although practice was continuous, time on target scores (to .001 min.) was obtained for each minute of practice by having two banks of clocks which scored for alternate minutes.

Experimental method.—Of the seven groups, six were experimental groups and one was a control group. All of these groups were given 32 min. of practice.

The control group practiced continuously for 32 min. The experimental groups were given either a 10-min. or 2-hr. rest after 4, 16, or 28 min. of prerest practice. The three groups that had the 2-hr. rest will be called A_4, A_{16}, and A_{28} depending on the minutes of practice after which the rest was given. Similarly, the three groups given a 10-min. rest will be labeled B_4, B_{16}, and B_{28}. These particular points in acquisition (4, 16, and 28 min. of prerest practice) were selected because the Two-Hand Coordination Test has four different trial patterns and there is some evidence that these patterns are not equally difficult (5). Thus all experimental groups were equated in terms of pre- and postrest trial difficulty.

Results

In the analysis of the results the scoring unit employed was mean time on target for the blocks of four *Ss* tested together.

Figure 1 gives the acquisition curves in terms of minutes on target for the seven groups. The curve for the control group is the average of all the groups that have not had the rest. The curves for the experimental groups are shown only after the imposition of the rest. In each case the mean of the experimental groups is given for the last prerest trial. It will be noted that after a brief period

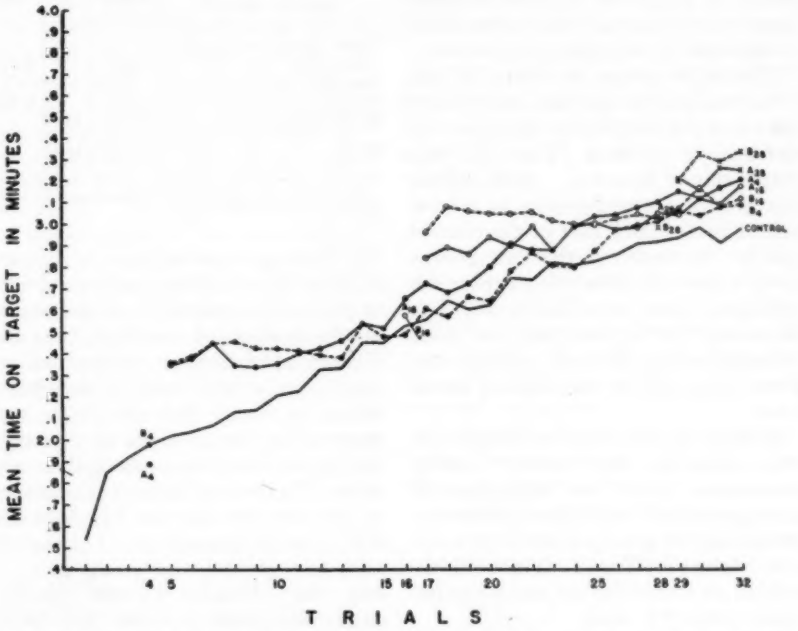


FIG. 1. Mean time on target for averaged control and experimental groups. Dotted lines (B) refer to 10-min. rest groups and solid lines (A) to 2-hr. rest groups.

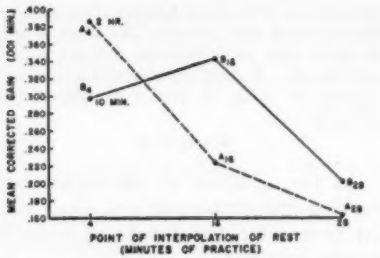


FIG. 2. Mean corrected gains over rest for 2-hr. (A) and 10-min. (B) rest groups for 4-, 16-, or 28-min. prerest practice

of negative acceleration the acquisition curve for the control group is linear. This is typical of performance curves for motor learning under these conditions. For all the experimental groups there is a marked increase in performance on the first postrest trial. This is followed by a period in which the acquisition curve seems to be essentially flat after which it becomes an increasing linear one.

The solid curves are those for the 2-hr. rest groups and the dotted ones are those for the 10-min. groups. At the end of practice (Trial 32) two points should be noted. First, all the rest groups are performing at a level that is better than that of the control group. Secondly, if the two groups given a rest just four trials earlier are excluded, there is an indication that the curves for the two rest conditions are approaching different asymptotes. These facts will be examined in detail later.

In order to find whether the groups were differing significantly among themselves before the imposition of the experimental variables, a between-groups/within-groups analysis of variance was run for Trial 4. This analysis yielded an F that did not reach significance at the 5% level.

Figure 2 shows the mean corrected gains over rest for the 2-hr. (A) and the 10-min. (B) rest groups for rests given after 4-, 16-, and 28-min. prerest practice. The absolute gains for the rest groups were corrected by subtracting from them the gains made by the control group on corresponding pairs of trials.

For all rest groups, the performance on the first postrest trial was superior to that of the control group on the corresponding trial. Table 1 gives the summary of the analysis of variance for the gains for the experimental groups. It will be seen that only the F for the point of interpolation of rest reaches significance. Comparison of group means by means of t tests using

TABLE 1

ANALYSIS OF VARIANCE FOR EXPERIMENTAL GROUPS' POSTREST GAINS

Source of Variation	Sum of Squares	*df*	*F**
Point of interpolation of rest (PI)	.6606	2	5.856
Length of rest (L)	.0260	1	<1
PI × L	.3714	2	2.406
Within groups	8.1185	144	
Total	9.9765	149	

* For *df* = 2 and 150, 3.06 = 5% level and 4.75 = 1% level.

the within-groups variance as the estimate of the population variance shows, as would be expected from the results of the analysis of variance, that the differences between rest groups are not significant at any point of interpolation of rest. For the 2-hr. rest groups, the group given rest after 4 min. prerest practice gains significantly more (5% level or better) than either of the other two groups, which do not differ among themselves. For the 10-min. rest groups, the group given a rest after 16-min. prerest practice gained significantly more (5% level) than the group given a rest after 28-min. prerest practice, but the other

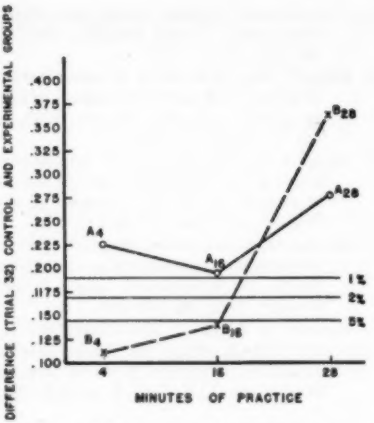


FIG. 3. Differences on Trial 32 between control and 2-hr. (A) and 10-min. (B) rest groups. Levels of significance of the differences are indicated by horizontal lines.

between-group differences are not significant.

As was mentioned above, at the end of practice the performance of the various groups has some interesting aspects. A between-groups/within-groups analysis of variance of the performance data for Trial 32 gave an F significant beyond the 1% level of confidence. The differences between the performance of the experimental groups and the control group on Trial 32 are presented in Fig. 3. The levels of significance of these differences are indicated by the three horizontal lines. Again the 2-hr. rest groups are labeled A and the 10-min. B. All of the 2-hr. rest groups were performing better than the control group but this was true only for the 10-min. rest group that had a rest just four trials previously.

DISCUSSION

In terms of these results certain general statements can be made within the limits of the present experiment. In the performance of a continuous two-hand motor task the introduction of a rest during acquisition leads to better postrest performance when such performance is compared with that of *S*s that have not had a rest. On the basis of the comparative gains of the various experimental groups the gain over rest, in terms of the immediate effect of the rest, does not seem to be a function of the length of rest. However, given continued practice after rest, the longer rest seems to have longer-lasting effect. At the end of practice all groups given the longer rest are performing at a level above that of the no-rest group, while this is true only for the short-rest group given rest late in practice.

The effects of rest in terms of gain are some decreasing function of the point in acquisition at which the rest is given. Although this function appears to be different for the two lengths of rest, the lack of a significant interaction between the point of interpolation of the rest and the length of rest and the lack of significant differences betweeen the rest groups at any given point of interpolation argue against this.

SUMMARY

The present study investigated the effects on performance of a continuous, two-hand motor task of the interpolation of rests of either 2 hr. or 10 min. after 4, 16, or 28 min. of prerest practice. The results may be summarized as follows:

1. Length of rest was not a differential determiner of performance in terms of postrest gains.

2. With continued practice, there was an indication that the longer rest was more beneficial than the shorter one.

3. The point of interpolation of rest affected performance significantly. In general, the shorter the prerest prac-

tice period, the greater was the gain over rest.

4. There was no significant interaction between the length of rest and the point of interpolation of rest.

5. For all points of interpolation and both lengths of rest the first post-rest trial performance was superior to that of a control group that had no rest.

(Received September 5, 1952)

References

1. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–410.
2. Bell, H. M. Rest pauses in motor learning as related to Snoddy's hypothesis of mental growth. *Psychol. Monogr.*, 1942, **54**, No. 1. (Whole No. 243).
3. Irion, A. L. Reminiscence in pursuit-rotor learning as a function of length of rest and of amount of pre-rest practice. *J. exp. Psychol.*, 1949, **39**, 492–499.
4. Kimble, G. A., & Horenstein, B. R. Reminiscence in motor learning as a function of length of interpolated rest. *J. exp. Psychol.*, 1948, **38**, 239–244.
5. Melton, A. W. (Ed.) *Apparatus tests*. Report No. 4, Army Air Force Aviation Psychology Program Research Reports. Washington: U. S. Govt. Print. Off., 1947.

Journal of Experimental Psychology
Vol. 45, No. 4, 1953

THE RELATION OF CONDITIONED RESPONSE STRENGTH TO ANXIETY IN NORMAL, NEUROTIC, AND PSYCHOTIC SUBJECTS[1]

KENNETH W. SPENCE AND JANET A. TAYLOR

State University of Iowa *Northwestern University*

Following the discovery and popularization of the classical conditioning technique by Pavlov and his co-workers, attempts were made by numerous investigators to relate performance in a conditioning situation to a variety of personality and intellectual variables. One particular area of investigation involved the comparison of the conditioning of normal individuals with specific types of neurotic and psychotic patients. These early experiments (**1, 11, 15, 17**) were either empirical in nature or were concerned with testing certain theories developed by Pavlov (**12, 13, 14**). In order to explain individual differences in conditioning observed among his experimental animals, Pavlov had postulated different types of nervous systems. He further related these neurological types to several neurotic and psychotic disorders in human *Ss*, thus offering both an interpretation of certain varieties of psychopathology and of possible differences in conditioning among such groups. In reviewing these early studies Hilgard and Marquis (**9**) concluded that on an empirical level no clear-cut relationships had been established, the relative performance of groups varying from experiment to experiment and from one type of conditioning situation to another, and that at best, the theories put forth by Pavlov were premature.

In recent years there appears to have been a revival of interest in the relationship between conditioning and personality variables, with emphasis on anxiety. A series of studies by Welch and his co-workers (**16, 22, 23**) have demonstrated that hospitalized adults and children diagnosed psychiatrically as exhibiting pathological anxiety condition more rapidly in a PGR situation than do normal individuals. These experiments, it will be noted, differ from the early studies in that a particular symptom pattern, pathological anxiety, has been related to conditioning independent of the diagnosis of the patients involved. The interest of previous experiments, on the other hand, had been in establishing differences in conditioning among the diagnostic groups themselves.

A second group of investigations have been concerned with the relationship of conditioning to degree of anxiety, variously defined, in normal individuals. Taylor (**20**), proceeding within the framework of Hullian theory, assumed first that response strength (R) in a conditioning situation is some positive function of total drive (D), and secondly that variations in what is described psychiatrically as "manifest" or overt anxiety reflect differences in drive level. Groups of anxious and nonanxious *Ss*, chosen on the basis of extreme scores on a test of manifest anxiety,[2]

[1] This investigation was supported by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service. The authors wish to acknowledge the very helpful cooperation of Dr. David Shakow and the members of the psychiatric staff of the Department of Psychiatry, Illinois Neuropsychiatric Institute. Acknowledgment is also made to Mr. George Faibish who served as *E* for the patient groups and to Miss Marguerite Craft who was *E* for a portion of the normal *Ss*.

[2] The manifest-anxiety scale developed for use in the Taylor study consisted of 65 items drawn from the Minnesota Multiphasic Personality Inventory judged by clinicians to be indicative of manifest anxiety. These items, along with 125 nonanxiety "buffer" statements from the MMPI, were administered in group form to introductory psychology students. Selection of *Ss* was made by taking the upper and lower extremes of the distribution. The test was subsequently revised, 15 of the original anxiety items being eliminated because of low correlation with total test scores in the original group tested, and the padding items being expanded in number. It is

were conditioned by means of the eyelid technique. The anxious *Ss* exhibited a significantly greater amount of conditioning, thus supporting the hypotheses being tested. These results were duplicated in a subsequent study by the writers (19), *Ss* receiving high scores on the manifest-anxiety scale again showing a conditioning performance superior to those with low scores.

In another investigation also using the manifest-anxiety scale Hilgard, Jones, and Kaplan (8) correlated scale scores with eyelid-conditioning performance obtained from an unselected group of *Ss*, and found a positive but statistically insignificant correlation coefficient. In a second part of the investigation involving conditioned discrimination, a significant correlation was obtained, *Ss* with high anxiety scores exhibiting more CR's to the negative stimulus. They explained these results on a very different theoretical basis than that used by the present writers.

A relationship between anxiety and simple conditioning was also established by Bitterman and Holtzman (2). Defining anxiety in terms of a composite rating obtained from a test battery and observational material, these *Es* found that *Ss* receiving high anxiety ratings (the upper 50% of their group) exhibited superior PGR conditioning when compared to those with low anxiety ratings (the lower half of the distribution).

The present study represents a further investigation of the influence of personality variables on conditioning, with particular reference to neurotic and psychotic groups. Primary interest was in comparing the performance of normal *Ss* with that of neurotic and psychotic *Ss*. The relation of anxiety (drive) level to conditioning performance in normals and patients was also studied.

Clinical Concepts of Anxiety and Their Implications for Conditioning

Beginning with the work of Freud (7) the clinical concept of anxiety has come to be widely accepted by clinicians as a key explanatory principle in the functional behavior disorders (3, 4, 10, 24). According to these clinical conceptions, pathological symptoms in both neuroses and psychoses are regarded either (*a*) as primarily direct, overt expressions of anxiety or (*b*) as a defense against anxiety. In the case of the former, certain observable symptoms (e.g., complaints of worry and fatigue, digestive upsets, etc.) are defined as manifest anxiety and assumed to be the accompaniment of an internal emotional state. In the second instance, the symptoms are said to develop as defenses against the hypothetical internal anxiety, and as such serve the function of reducing or avoiding this inner anxiety state. Manifest anxiety is, then, directly observed but the presence of internal anxiety is inferred from the appearance of the "defensive" reactions.

The studies of conditioning previously mentioned (2, 8, 19, 20) have investigated the relationship of manifest or overt anxiety to conditioning. It will be recalled that in the investigations carried out by the writers (19, 20) the specific assumption was made that Hull's intervening variable (*D*) is some positive function of manifest anxiety (as measured by the anxiety scale) and that, therefore, higher degrees of anxiety (drive) would be accompanied by a higher level of conditioning performance.

The relationship between Hull's *D* and the clinical concept of the internal anxiety postulated to lie behind the defenses is not clear. However, clinicians appear to have assumed that the functional properties of the state are the same as those underlying manifest-anxiety symptoms. On this basis one would be led to predict that different degrees of the hypothetical internal anxiety inferred from these pathological symptoms might also be expected to be accompanied by different levels of drive.

If the theoretical notions outlined above are correct, psychotic and neu-

this form of the scale that has been used by the writers in their studies. Detailed information concerning the various forms of the scale is reported elsewhere (21).

rotic individuals would be expected, in general, to exhibit a higher level of performance in conditioning than would normal *Ss*, because of the greater frequency of manifest-anxiety symptoms with their accompanying internal emotional responses and of the greater intensity of the hypothetical anxiety underlying other types of pathological behavior.[3] The plan of the present experiment, then, is to compare the conditioning rates of neurotic, psychotic, and normal *Ss* as an experimental test of these hypotheses.

Procedure

Normal subjects.—Subjects constituting the normal group were volunteers from night classes in introductory psychology at Northwestern University College. The classes of which these *Ss* were members had been given the revised form of the manifest-anxiety scale previously described so that anxiety scores were available for each of them. In addition, some personal data (age, sex, occupation, schooling, previous or present psychiatric treatment) were obtained. None of the normal *Ss* run in the conditioning experiment had had, or admitted to, any psychiatric attention.

Sixty normal *Ss* were run in the experiment, of which ten were excluded because of apparatus failure or *E*'s error. Of the remaining 50 *Ss* the records of five were discarded because they did not meet certain latency criteria established as a method of eliminating *Ss* believed to be making voluntary responses to the CS. These criteria will be discussed in a later section. The final group of normal *Ss* whose records were retained thus numbered 45.

Patient subjects.—The neurotic and psychotic *Ss* were obtained from the regular psychiatric in-patient and out-patient services of the Illinois Neuropsychiatric Institute. Some screening of patients used in the experiment was involved in that the psychiatrist in charge of each case referred only those individuals who were deemed capable of cooperating in the experiment and who would not be sufficiently upset emotionally by the experience as to retard the progress of therapy.

Psychiatric diagnoses were obtained for the patient group from the individual files maintained by the institution. Several additional kinds of information were also gathered for each patient: a psychiatric checklist[4] which was filled out for each patient by his psychiatrist, items of personal data drawn from the case history, and several psychological tests including an individual form of the manifest-anxiety scale and the Rorschach. The latter were usually administered several days after the conditioning session.

The patient group was more heterogeneous with respect to age, education, and occupational status than the normals, but as experimental evidence (9) fails to indicate that these variables are influential in conditioning, no importance was attached to these differences.

Seventy-two patients were run in the conditioning situation. The records of 17 of these were excluded from the final results for the following reasons: 4 for equipment failure, 3 owing to inability or unwillingness to cooperate throughout the conditioning session, and 10 because they did not meet the latency criteria set up to exclude voluntary responders. Of the 55 *Ss* remaining, 34 were diagnosed as neurotic and 21 as psychotic.

Apparatus and method of recording.—The normal and patient groups were tested in different locations but as the two sets of apparatus were duplicates and the physical arrangements similar, the same description of the apparatus holds for both groups. The *S* was seated in a straight-backed chair in a semidarkened room containing the conditioning apparatus. He was instructed to blink upon receiving a verbal ready signal and then to fixate on a 6-cm. circular milk glass window or disc, placed at a distance 6 ft. in front of him. The brightness of the disc between trials was 1.2 apparent ft.-candles.

The CS consisted in an increase in brightness of the disc to 23.8 apparent ft.-candles. The duration of the CS on each trial was 520 msec.

[3] Clinicians have offered specific hypotheses concerning the relationship of the hypothetical internal anxiety to particular syndromes (e.g., paranoia, hysteria) which are concerned with the effectiveness of the specific symptom pattern in reducing or avoiding anxiety. Thus, if these assumptions are correct, differences in conditioning would also be expected among diagnostic groups as a function of the postulated effectiveness of the defense. A further study by the present writers is now being undertaken in order to test some of these clinical speculations.

[4] The psychiatric checklist was developed for use in the present investigation by Mr. George Faibish. This checklist consisted of common behavior patterns found in the functional behavior disorders and included a detailed list of anxiety symptoms. The psychiatrist was asked to indicate absence or severity of each symptom and further to rate the patient as to degree of overt or manifest anxiety exhibited and the amount of internal, underlying anxiety judged to be present.

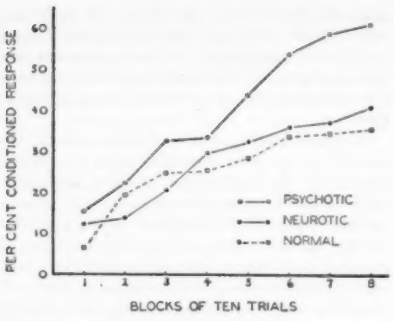


FIG. 1. Conditioning curves for three groups of *Ss*

Following the onset of the CS by 470 msec., the UCS, a puff of air, was administered in the region of the external canthus of the right eye. The air puff was produced by the fall of a 30-mm. column of mercury, magnetically released, in a manometer. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec., and fixed according to a prearranged schedule, were used. The onset of the CS was controlled by the closing of a switch on an electronic timer. The latter controlled the duration of the CS and the onset of the UCS at the appropriate time interval. The onset of the stimuli was recorded on a polygraph by means of an electromagnetic marker.

Eyelid movement was recorded by means of an electromagnetic system devised in the Iowa laboratory (**18**) that provided a linear record of the position of the eyelid at all times. Closure of the lid was recorded on the moving polygraph paper as a deviation downward from a horizontal line, the extent of this deviation being a linear function of the amplitude of eyelid movement. A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval 200 to 470 msec. following the onset of the CS. Responses with a latency of less than 200 msec. were classified as original responses (alpha or beta) to the CS and were not included in the data.

Studies (**18, 19**) in the Iowa laboratory had shown that some *Ss* seek to facilitate conditioning by voluntarily closing their eye to the conditioned light signal and keeping it closed until after the occurrence of the UCS. The patterns of such responses are characteristic in form and tend to be short in latency. In order to eliminate such *Ss* from the present experiment objective latency criteria utilized and described in detail in a previous study by the writers (**19**) were employed. Briefly, the data for all *Ss* who gave more than 50% CR's with latencies less than 300 msec. were excluded from further consideration.

Conditioning procedure.—Each *S* first received three presentations of the light alone in order to check for any initial conditioned response tendencies to the light. A single puff of air, unaccompanied by the light, was then administered and spontaneous blinking recorded for 40 sec. following onset of the puff. Immediately following the preliminary trials, all *Ss* were given 80 paired presentations of the light and air puff. After the final conditioning trial, ten extinction trials were administered.

At the end of the session the normal *Ss* were questioned as to their understanding of the purpose of the experiment and warned not to discuss the experiment with other members of the class.

RESULTS AND DISCUSSION

Amount of conditioning.—Figure 1 presents curves of acquisition of the CR, based on the per cent of anticipatory responses occurring in successive ten-trial blocks, for the three experimental groups. As can be seen by inspection of the curves, the psychotic group is above both the neurotic and normal *Ss* in every block of trials. In contrast, there is some crossing of the curves for the neurotic and normal groups, although the neurotics are consistently superior to the normals in the last five of the eight blocks of trials.

A further comparison of the three groups is contained in Table 1, which gives the total amount of conditioning exhibited by each group. The psychotic *Ss*, it will be seen, are superior to the other groups while the neurotics are only slightly above the normals.

TABLE 1

MEAN PERCENTAGE OF CR'S EXHIBITED ON 80 CONDITIONING TRIALS BY THE THREE EXPERIMENTAL GROUPS

Group	*N*	Mean	*SD*	σ_M
Normal	45	26.05	18.41	2.78
Psychotic	21	40.41	26.05	5.83
Neurotic	34	28.57	24.54	4.28

TABLE 2

ANALYSIS OF VARIANCE FOR TRIALS 1–80

Source	*df*	Mean Square	*F*
Between groups	2	121.95	2.40
Between *Ss*	97	50.88	
Between trials	7	137.50	48.2*
Trials × Groups	14	6.96	2.44*
Pooled *Ss* × Trials	679	2.85	
Total	799		

* Significant at the .01 level of confidence.

In order to determine the significance of the obtained differences, an analysis of variance for repeated measures described by Edwards (6) was performed. The results of this analysis are summarized in Table 2. The between-groups *F* was found to approach, but not to reach, the value necessary for significance at the .05 level of confidence. This result suggested that the one extreme group (psychotics) might be significantly different from one or both of the remaining groups. Accordingly, a *q* statistic, reported by Dixon and Massey (5) to be appropriate for such situations, was computed to test the hypothesis that there was no difference between the psychotics and the two remaining groups. The results of this test indicated that the null hypothesis could be rejected at the .05 level of confidence.

Further inspection of the results of the analysis of variance shown in Table 2 reveals that the *F* for interaction between trials and groups was significant beyond the .01 level of confidence. This appears to indicate that the slopes of the conditioning curves for the three groups were not parallel.

A breakdown of the neurotic group in terms of specific diagnosis revealed that *Ss* in one category, anxiety states, were superior in amount of conditioning when compared to the remaining subgroups. The total amount of conditioning for the 9 *Ss* with the diagnosis of anxiety state was 39.17% as opposed to 26.90% for the other 23 neurotics and 26.05% for the normal group. A *t* test, employed to determine the significance of the difference between anxiety states and normals, yielded a value of 1.47, which is not significant at the .05 level of confidence for a single-tailed hypothesis.

The results of the comparison of the three experimental groups indicate that the hypothesis originally offered was only partially confirmed, only the psychotic *Ss* exhibiting significantly higher conditioning than normals. Furthermore, the slight, statistically insignificant superiority of neurotics over the normal group was primarily due to the high conditioning scores of one diagnostic group, anxiety states. Although the superiority of the anxiety neurotics in whom manifest- or overt-anxiety reactions constituted the predominant symptomatology was not significantly different from the normals at a high level of confidence, there is a suggestion that a significant difference might have been obtained had a larger number of anxiety neurotics been available.

The fact that anxiety neurotics tended to be superior to normals while *Ss* falling into other neurotic groups were not has an interesting implication for clinical theory. As was discussed previously, clinicians have described many pathological symptoms, particularly those of the neurotic type, as "defenses" against an internal anxiety state, serving to reduce or avoid this state. The failure to find differences between the conditioning of normals and neurotic *Ss* whose symptoms consist primarily of defense reactions gives some support to these speculations.

Conditioning as related to scores on the Manifest Anxiety scale.—Previous

TABLE 3

MEAN PERCENTAGE OF CONDITIONING EXHIBITED IN 80 TRIALS BY THE FOUR NORMAL SUBGROUPS

Mean Anxiety Score for Group	*N*	Mean	*SD*
5.25	12	15.31	21.15
9.30	10	21.75	20.01
15.22	9	24.45	15.40
26.83	12	38.12	25.01

studies (**19, 20**) have revealed that there is a relationship between level of conditioning performance and extreme scores on the manifest-anxiety scale. Since anxiety scores were available for the three groups in the present experiment, it was possible (*a*) to determine the correlation between anxiety scores and conditioning in normal *Ss* throughout the entire range of scores, and (*b*) to discover whether differences in anxiety scores are also accompanied by differences in level of responding in the patient groups.

The normal *Ss*, all except two of whom had taken the anxiety scale, were divided into four approximately equal subgroups according to anxiety scores. Exact equality of the number in each subgroup was not possible to obtain and still have a specific score appear in only one group. The mean conditioning scores for these four subgroups are presented in Table 3 in which it may be seen that level of conditioning becomes greater as anxiety scores increase. The results of a *t* test indicate that the difference in conditioning between the two extreme groups was significant at the .01 level of confidence for a single-tailed hypothesis. When the two lower groups and the two upper groups were combined so that approximately the upper and lower 50% of the distribution were being compared, the *t* test yielded a value significant at approximately the .02 level, again for a single-tailed hypothesis.

In order to determine the nature of the function holding between conditioning performance and anxiety scores, total amount of conditioning has been plotted in Fig. 2 against mean anxiety score for each of the subgroups. The relation, as may be seen from the figure, appears to be linear. This finding differs somewhat from that obtained in two unpublished experiments conducted in the Iowa laboratory. The latter suggest that the relation is curvilinear with a greater difference occurring between the highly anxious *Ss* and those at the median than between the nonanxious *Ss* and those at the median.

The distribution of anxiety scores for the patients was very different from that found with the normal *Ss*, having a mean of about 32 and a marked negative skew, in contrast to a slight positive skew and a mean of about 14 for the normals. Since the range of anxiety scores for the patient group was so constricted, it did not seem reasonable to follow the same procedures as those used with the normal *Ss* to determine the relationship between conditioning and anxiety scores. Instead it was decided to compare *Ss* with extreme anxiety scores in terms of the patients' distri-

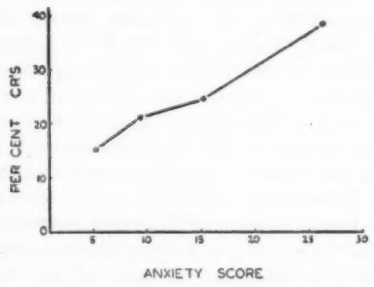


FIG. 2. Percentage of CR's in 80 conditioning trials as a function of anxiety level

bution. Selection of *Ss* for this purpose was made by taking those scoring in the lower 26% (scores of 22 or less) and the upper 27% (scores of 43 or more) of the distribution.

The mean amount of conditioning exhibited by the low-scoring group was 30.00% as compared to 39.84% for the high-scoring *Ss*. Although the difference in conditioning between the two groups was in the expected direction, the *t* computed to test its significance was only .92.

Implications for Diagnosis

A number of studies (**2, 16, 19, 20, 22, 23**), including the present one, have now demonstrated that groups with high manifest anxiety, variously defined, exhibit a significantly higher level of conditioning performance than do *Ss* with low anxiety. This has led some to suggest that conditioning is a potential diagnostic tool. Bitterman and Holtzman (**2**), for example, stated on the basis of their study which showed a significant difference between the upper and lower 50% of their group that "conditioning techniques may be of considerable value in . . . psychiatric screening . . ." (**2**, p. 13).

It should be pointed out, however, that such a hope, based as it is on the demonstration of significant differences in conditioning between extreme groups, is at best premature, since such a difference does not necessarily indicate the presence of a high correlation coefficient between the two variables.

In the present experiment, for example, in which a highly significant difference in conditioning was found between *Ss* at the extremes of the anxiety scale, the product-moment *r* between the two variables was nevertheless only .25. While significantly different from zero at the .05 level of confidence, computation of the standard error of estimate indicated that this magnitude of coefficient would yield only a 3% improvement in prediction over guessing.

It does not appear likely, moreover, that the relatively low correlation coefficient between conditioning and manifest-anxiety scores obtained in the present study is confined to the particular definition of anxiety employed here. In studies employing other criteria of anxiety, as well as in the comparison of anxiety neurotics and normals in the present investigation, the variances of conditioning scores, when reported, indicate considerable overlapping between groups and suggest that the magnitude of the correlation coefficient would not be high. Thus it would appear that anxiety is only a relatively minor variable contributing to individual differences in conditioning. While the conditioning technique seems to be valuable in determining differences in anxiety between extreme groups, it does not at the moment appear useful for individual diagnostic purposes.

Summary

The present investigation was designed to investigate differences in conditioning performance between normal *Ss* and individuals diagnosed as neurotic or psychotic. Forty-five normal *Ss* were obtained from night school psychology classes while 34 neurotic and 21 psychotic *Ss* were drawn from the in- and out-patient services of a neuropsychiatric hospital. All groups were run in a conditioned-eyelid situation under the same experimental conditions.

The results indicated that the psychotic group was significantly higher in conditioning performance than either the neurotics or normals. The neurotic group, on the other hand, showed only a slight, statistically in-

significant, superiority over the normals. Conditioning performance was also related to scores on a test of manifest anxiety for both the normal and patient groups. Patient *Ss* with extremely high anxiety scores (in terms of the distribution for the combined neurotic and psychotic groups) exhibited a higher performance than did patients with low scores but the difference was not statistically significant.

For the normal group, *Ss* with extreme scores on the anxiety scale differed significantly in conditioning performance. The correlation coefficient computed to determine the magnitude of the relationship between the two variables however was only .25.

The results were discussed in terms of their possible implications for clinical theory concerning the relationship between anxiety and the role of pathological symptoms and for the potentialities of the conditioning technique as a diagnostic tool.

(Received August 29, 1952)

References

1. Bender, L., & Schilder, P. Unconditioned and conditioned reactions to pain in schizophrenia. *Amer. J. Psychiat.*, 1930, **87**, 365–384.
2. Bitterman, M. E., & Holtzman, W. H. Development of psychiatric screening of flying personnel: III. Conditioning and extinction of the galvanic skin response in relation to clinical evidence of anxiety. U.S.A.F. School of Aviation Medicine, Project No. 21–37–002.
3. Brown, J. F. *Psychodynamics of abnormal behavior.* New York: McGraw-Hill, 1940.
4. Cameron, N. *The psychology of behavior disorders: A biosocial interpretation.* Boston: Houghton Mifflin, 1947.
5. Dixon, W. J., & Massey, F. S., Jr. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.
6. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
7. Freud, S. *The problem of anxiety.* New York: W. W. Norton, 1936.
8. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–100.
9. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
10. May, R. *The meaning of anxiety.* New York: Ronald, 1949.
11. Mays, L. L. Studies of catatonia: V. Perseverational tendencies in catatonic patients. *Psychiat. Quart.*, 1934, **8**, 728–35.7
12. Pavlov, I. P. *Lectures on conditioned reflexes.* (Trans. by W. H. Gantt) New York: International, 1928.
13. Pavlov, I. P. Neuroses in man and animals. *J. Amer. med. Ass.*, 1932, **99**, 1012–1013.
14. Pavlov, I. P. An attempt at a physiological interpretation of obsessional neurosis and paranoia. *J. ment. Sci.*, 1934, **80**, 87–197.
15. Pfaffman, C., & Schlosberg, H. The conditioned knee jerk in psychotic and normal individuals. *J. psychol.*, 1936, **1**, 201–205.
16. Schiff, E., Dougan, C., & Welch, L. The conditioned PGR and the EEG as indicators of anxiety. *J. abnorm. soc. Psychol.*, 1949, **44**, 549–552.
17. Shipley, W. C. Studies of catatonia. VI. Further investigation of the perseverational tendency. *Psychiat. Quart.*, 1934, **8**, 736–744.
18. Spence, K. W. Learning and performance in eyelid conditioning as a function of the intensity of the UCS. *J. exp. Psychol.* 1953, **45**, 57–63.
19. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
20. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
21. Taylor, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
22. Welch, L., & Kubis, J. Conditioned PGR (Psychogalvanic Response) in states of pathological anxiety. *J. nerv. ment. Dis.*, 1947, **105**, 372–381.
23. Welch, L., & Kubis, J. The effect of anxiety on the conditioning rate and stability of the PGR. *J. Psychol.*, 1947, **23**, 83–91.
24. White, R. W. *The abnormal personality.* New York: Ronald, 1948.

QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Founded in 1948

Published Quarterly in February, May, August, and November

for the

EXPERIMENTAL PSYCHOLOGY GROUP

Editor: D. Russell Davis, University of Cambridge

Subscriptions from the United States should be entered now for the next year. 30 shillings per year (current rate $4.20).

Order from

W. Heffer & Sons, Ltd.
Cambridge, England

WILEY BOOKS

Man's Senses

Treated in a modern manner

THE HUMAN SENSES

By FRANK A. GELDARD

Professor of Psychology, University of Virginia

Here is the first, modern, comprehensive description of man's senses viewed from a psychophysiological point of view. The author examines each of man's senses in turn in a double perspective—first through the physics of stimuli, and then through the anatomy and physiology of the sense organs.

This new book pinpoints the place of sensation in psychological and physiological knowledge. Unfolding the drama of each of the senses, it stresses their importance in the total behavior picture. For, as the author says in the Preface, " . . . the high road to an understanding of human nature is through an appreciation of the role played by man's senses in the attainment of knowledge and the regulation of behavior."

Of particular value to psychologists are the chapters on the often neglected cutaneous senses. Welcome too, will be the emphasis on the importance of sensation in modern engineering psychology and recent human engineering developments. Enhancing the presentation, over 100 illustrations highlight the reading material and graphically bring home the points discussed.

As a broad, unified, up-to-date presentation of the mainsprings of human action, *The Human Senses* belongs in the library of every psychologist.

1953 **365 pages** **$5.00**

Send today for an on-approval copy

JOHN WILEY & SONS, Inc.

440 Fourth Ave. **New York 16, N. Y.**